

Tough talk on surrogate birth

Is it wholly wrong that married couples unable to conceive children naturally should pay others to perform that service? Mistakenly, Britain is behaving as if there were no doubts.

It was always on the cards that the British government, by its failure to legislate more quickly on the recommendations of the Warnock committee, would allow a muddle to develop. It seems to have a useful ally in that cause in the British Parliament. The most urgent need is that the position of those wishing to make observations of living embryos should be clarified; the Warnock committee advocated experiments under licence with embryos up to 14 days, whereafter any experiment would become the cause of a criminal prosecution. It would be good to think that the government has recognized the unwisdom of that proposal, and that it plans to legislate for licensing, coupled with supervision by independent committees, allowing researchers more flexibility, but there is no way of telling. Meanwhile, research limps along uncertainly, under the non-statutory supervision of a committee established by the Medical Research Council and the Royal College of Obstetricians and Gynaecologists — and the Houses of Parliament chip away piecemeal at what should be the government's responsibility. The dangers of this approach are well illustrated by the fate of last year's instant legislation, begun in the House of Commons, when the sale of surrogacy services became an offence. One consequence appears to have been that British couples who are infertile are seeking the service overseas, which the House of Lords will this week attempt to ban by law, making it illegal for agencies in Britain to offer to provide such services even with the help of surrogate mothers elsewhere. Presumably the result will be that the agencies concerned will operate offshore.

This is the inevitable result of attempts to prevent well-intentioned people from using what seems to them a natural way of meeting a natural need. Procreative instincts have an adaptive significance for all species and, while they may become a nuisance in long-lived human societies, cannot be repressed by legislation. Moreover, it is natural that couples should prefer genetically related to unrelated children: Dawkins's concept of the selfish gene, not to mention a great deal of sociobiology, refers. The Warnock committee, after eighteen months of careful study, was itself divided on the issue; the majority would have made surrogacy a criminal offence when money changes hands, partly by analogy with the law covering the adoption of children in Britain. But a minority wisely argued that circumstances might arise in which paid surrogacy is justified. (It seems to be common ground that the willingness of some women to bear children for close relatives, another illustration of the selfishness of genes, should not be prevented by law if no money changes hands.) What, for example, if a woman is infertile as a result of a hysterectomy? This is one of the questions raised last weekend by Dr Robert Edwards and Mr Patrick Steptoe, the pioneers of *in vitro* fertilization in Britain. It is hard to deny their case for surrogacy, ideally based on a fertilized egg from the infertile woman, but the British Parliament would even stop that.

The weakness in the British Parliament's position on surrogacy is that it confuses two quite unrelated emotional issues. Some people believe simply that unfamiliar practices are wicked, and should be stopped. This is, after all, the Parliament that required that men with red flags should walk in front of the

early motorcars. The other cause for anxiety, legitimate as it happens, is that surrogacy might become a vehicle for the exploitation of those concerned, either surrogate mothers or the putative genetic parents. So what the House of Lords should have been doing this week, and what the House of Commons will in due course have to decide, is how to protect the vulnerable parties to these transactions without denying remedies to those determined to have children.

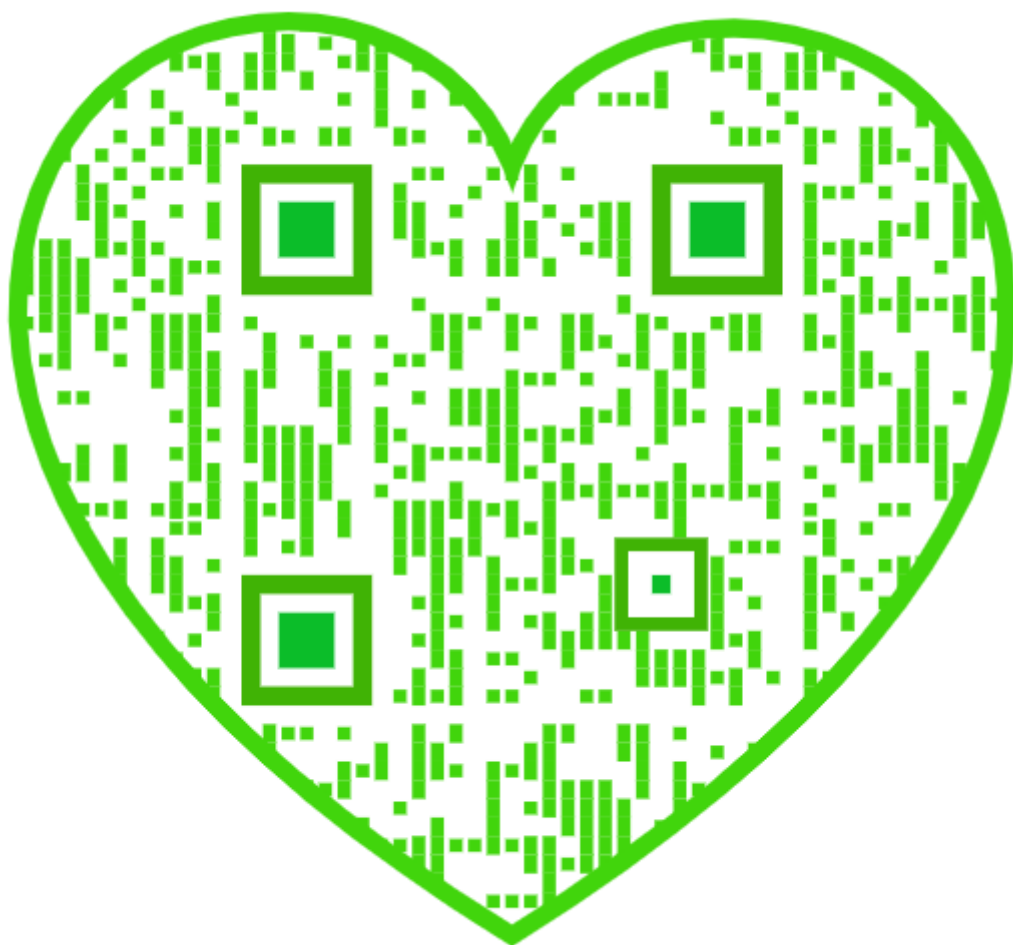
The best solution, for which there is already a good working model, is that arrangements for surrogacy should be made not by commercial agencies but by corporations with the legal status of charities (which, in Britain, is what adoption societies are). But since it is wrong to hazard the lives of women by childbirth on behalf of others, it is quite wrong to insist that the surrogate mothers should not be paid. On the contrary, they should be insured financially against risk, while compensation for the considerable time and trouble is the opposite of outrageous. Further, all those concerned in a surrogacy arrangement should have the benefit of independent counselling, of a kind unlikely to be fully convincing if provided by an agency with a profit to make. That constructive course, not futile repression, is the wise course to follow. □

Future for the birds

The future of the US shuttle is still in doubt, but the causes of what went wrong are clear.

THE reappointment of Dr James C. Fletcher as administrator of the US National Aeronautics and Space Administration (NASA) is one step towards the agency's rehabilitation (see p.100), but there is a long way to go. Although the technical cause of the failure of *Challenger* two months ago has not been finally pinpointed by the US President's commission of inquiry, it stands out a mile that the fault within the agency was simply the failure of bad news about the risk of rocket failure to travel upwards to those most in need of it, the people responsible for deciding how quickly to push ahead with the shuttle programme. No doubt before the commission's inquiry is over, it will become clear that the chief inhibition was the familiar phenomenon that the bearers of bad news usually win less than meagre credit. The shuttle pilots were probably right to have complained, earlier this week, that the shuttle should have been grounded at least since last summer.

Dr Fletcher's immediate problem is technical, how to get the shuttle safely off the ground again. He will have to say quite openly that he will not be able to do that on a shoestring, which may make him the bearer of bad news in the year when the Gramm-Rudman act comes into effect. He will also need to develop, or otherwise procure, an unmanned satellite launcher that can be used for the routine business of putting routine satellites into routine orbits, when it is pointless to risk people's lives. Meanwhile, the ambitious scheme for building a permanent space station should be postponed at least until the shuttle is made safe. But further ahead, he must find a way of encouraging constructive fault-finding within the agency. □



AIDS

Pasteur plans to pursue patent suit on virus

THE US Justice Department has told the Institut Pasteur in Paris, where virologist Luc Montagnier discovered the virus that causes acquired immune deficiency syndrome (AIDS), that "there is no case to answer" over the laboratory's complaint that the US National Institutes of Health (NIH) pirated Montagnier's discovery — and now profits from it by the sale of AIDS blood tests.

The French laboratory has, however, discovered what could be damning new evidence against NIH, and is now considering placing it before the Washington judge investigating whether NIH used the French-discovered virus for commercial gain, in breach of contract between NIH and the Pasteur.

Nobody disputes that Montagnier sent a sample of serum containing the AIDS virus, which he called LAV, to NIH virologist Robert Gallo in September 1983, but NIH claim that Gallo was unable to propagate LAV, and that he subsequently and independently isolated the virus he called HTLV-III. This and a blood test based upon it were rapidly patented and approved for use. HTLV-III tests are now marketed by five US companies, while the Pasteur's test (Elavia) based on Montagnier's virus is still awaiting US patent approval. This is despite the fact that the Pasteur applied for a patent for LAV and its associated test well before NIH applied for a similar patent for HTLV-III.

Pasteur sources claimed last week that the institute has new evidence, in the form of a letter describing research results, that Gallo's laboratory was making electron micrographs of the LAV virus in December 1983, three months after it was first received in a sample from the Pasteur. If this is confirmed, it would seem to contradict the NIH position that Gallo was unable to propagate LAV, and would strengthen at least the circumstantial evidence that HTLV-III is indeed derived from the French isolate. (HTLV-III and LAV have almost identical genetic sequences, but NIH have argued that this is merely because both viruses came from the same population of patients in New York.) According to the Pasteur, the letter in question describes December 1983 micrographs as pictures of LAV which indicate that it might be a form of lentivirus.

The electron micrographs in question were taken at the National Cancer Institute's Frederick Cancer Research Center. But Gallo and his colleagues scoff at the idea this constitutes evidence of skuldug-

gery. Gallo says the letter, along with all other "evidence", was supplied by his own laboratory to Montagnier. He achieved transient growth (evidenced by the electron microscope observations and reverse transcriptase activity) of the LAV sample supplied by Montagnier in a HUD-78 cell line, but for one week only and in small quantity: subsequent tests showed no evidence of viral replication. Attempts to grow LAV in the high-yielding H9 cell line were totally unsuccessful.

As for the significance of the timing, Dr Peter Fischinger of the National Cancer Institute points out that virus can be frozen and re-examined later, so the interval between the sample arriving and the date of the electron micrograph means nothing. Fischinger remarked that it would be surprising if Gallo had not made micrographs of his attempts to culture the virus. And on the charge that Gallo's isolate was in some way derived from LAV, Fischinger referred to recent results obtained using cytotoxic antibodies that kill cells infected with HTLV-III but not those infected with LAV or other isolates and so establish conclusively a real biological difference between the isolates.

In a further and ironic twist to the tale, the US Food and Drug Administration (FDA) last month approved the Pasteur's Elavia test as safe to use, but because of the lack of a patent, Pasteur sources are now concerned at the possibility of legal action against them if Genetic Systems Inc., the company due to market Elavia in the United States, actually sells any tests. According to the Pasteur, while Elavia lacks an independent patent, the five companies selling tests based on HTLV-III — and paying royalties under licence to NIH — could complain that Elavia was a "counterfeit" of the NIH test designed to avoid royalty payments, even though the French say the real case is the reverse.

Meanwhile, Elavia is said to be performing very well in tests at the British Public Health Laboratory in London, with results close to those of the Dutch (Organon) and British (Wellcome) tests approved for use "with flying colours" a few months ago. This may leave Elavia with some room on the British market, but in the United States the primary market in blood banks and hospitals is now thought to be effectively tied up by the five companies whose tests are now approved there. The French test shows fewer "false positives" (false indications that a patient has AIDS antibody) than most of the US tests, thus avoiding expensive follow-up

tests, but this may not now be enough to win a market share.

British scientists also have a small interest in Elavia, but their work also threatens to displace whatever toehold the French test may have in the United States. Elavia's technical advantage can be traced to the virus cultivated in a line of white blood cells, the CCRF-CEM line (or CEM for short) of T cells, that does not express certain antigens (HLA class two antigens) that can interfere with the results of blood tests.

The use of the CEM line was developed in London's Institute of Cancer Research. The director, Robin Weiss, says that "we have no claim on the CEM line", but his laboratory does have an agreement with the Pasteur that half of any royalties accruing to the Pasteur from use of the CEM line to prepare AIDS tests would return to Britain.

The Pasteur has made further developments "but we would still expect a proportion of the royalties" Weiss said last week, though "our goal has not been to make money but to develop a reliable test for public health". Weiss's laboratory has however developed its own more rapid test, using a British viral isolate and a completely different (and shorter) analysis technique, "competition assay", distinct from the enzyme-linked immunosorbent assay (ELISA) used by all other tests. This test is the one now approved and marketed by Wellcome in Britain. And according to Weiss, "we have not had one false positive in a million tests". Moreover, since the competition assay is complementary to other techniques, Weiss considers it a good follow-up test to check the results of others.

The conclusion is that whatever the result of the Pasteur's legal action in the United States, any mopping-up of the US market may go not to Genetic Systems and Elavia but to what are now more advanced tests such as Wellcome's, or to second-generation tests, now under development, based on the production of pure AIDS virus antigens by recombinant DNA techniques. It now seems clear that if the Pasteur wants to profit at all from tests for AIDS in the United States, it will have to prove that HTLV-III is derived from that sample of LAV that Montagnier sent Gallo back in 1983, and aim for a fraction of royalties on tests now on the market and derived from HTLV-III.

Robert Walgate & Tim Beardsley

Halley's comet data

By agreement with the European, Japanese and Soviet groups, *Nature* will publish on 15 May the first formal accounts of the data gathered by the four encounters with Halley's comet, together with related US observations. □

Halley's comet

Vega spacecraft reveal more dust than expected

Moscow

THE successful flight of the two Soviet space probes past the nucleus of Halley's comet showed that even from a distance of less than 8,000 km the nucleus of the comet was mostly obscured by dust. Estimates of the quantity of water being evaporated from the comet last week, a month after perihelion, range up to 2×10^{30} molecules per second. The three separate dust recorders in the first Vega package have shown that the mass of dust being lost is roughly a sixth that of the water. In round numbers, the water loss on 6 March was 30 tonnes an hour, that of dust roughly 5 tonnes an hour.

One practical consequence is that images from the two cameras on Vega 1 were unable to distinguish the nucleus of the comet within its dust cocoon. By the time of the encounter by Vega 2 on 10 March, the quantity of dust had fallen by a factor of two, for reasons not necessarily connected with the increasing distance of the comet from the Sun. But the camera images on this occasion are not as easily interpreted because of a pointing error, corrected only 10 minutes before the encounter. Even so, the Hungarian team working on the processing of images says it has extracted the information that the nucleus has a dumbbell shape and is less than 10 km across at the maximum.

The dust measurements have provided most of the surprises. Both Vega craft encountered the main dust sheet at about 100,000 km from the nucleus of the comet, in line with the expectation that the dust is responsible for the coma caused by the scattering of solar radiation. But some of the experimenters have been surprised to find recordings of dust particles from the comet as far away as 150,000 km.

Others say it is surprising that the dust recorders show the appearance of fine dust (10^{-15} grams or smaller) before that of more substantial grains. On the present view of the formation of the coma and the cometary tail, dust is carried from the solar side of the nucleus by the stream of evaporated water molecules until driven away from the Sun by radiation pressure, which leads to the prediction that larger grains should form the outermost sheath of the coma and the tail.

The sheer size of the dust particles has also been surprising. The Soviet originators of the photon experiment, an ingenious device by which the holes caused by dust particles penetrating the thin outer shell of the Vega spacecraft are measured optically and the accompanying plasma jet is detected electronically in such a way as to provide both the energy and momen-

tum of the incident particle, say the largest particle recorded by Vega 2 had a mass of 2×10^{-5} grams. It left a hole in the outer skin nearly half a centimetre across.

The approach to the comet by both spacecraft has also provided a wealth of information about the transition between the region dominated by the solar wind and that in which ions from the comet are predominant. At this stage, it is not clear

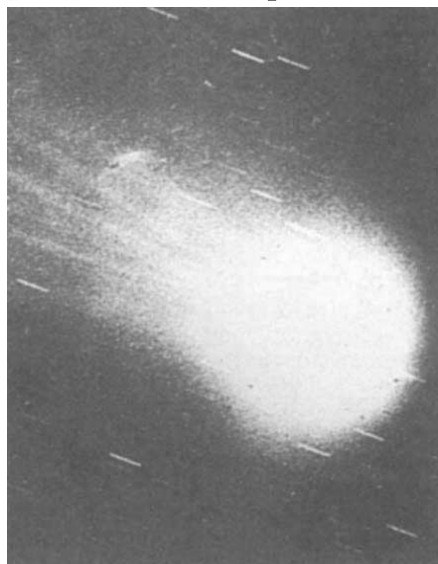


Image of comet Halley on 17 December 1985, taken with a 1-m reflector at the Alma-Ata Observatory by K.I. Churyumurov and Rspayev. North is up and East to the left.

whether either of the spacecraft detected the expected bow shock, representing the interaction between the solar wind and the cometary environment. Vega 1 may have passed through the bow shock before the instruments were switched on just over three hours before the encounter on 6 March; alternatively, the transition region may be more diffuse than expected.

At the second encounter on 10 March, both the intensity of the solar wind and the concentration of gas and dust from the comet were reduced by a half. The plasma detectors on Vega 2 have clearly shown how, on the inward passage towards the comet, ions (mainly protons) characteristic of the solar wind are gradually replaced by heavier ions, chiefly those of water and its photolysis products, between 150,000 km and 50,000 km from the nucleus.

Magnetic field measurements in the hours before the two encounters, essentially with similar results, also show the steady transition from the interplanetary to the cometary environment. On both occasions, the magnetic field increased (but with an apparent periodicity which is not yet explained) from the

12–15 nT (1 nT = 10^{-4} gauss) characteristic of the interplanetary solar wind regime to 75 nT within the cometary environment.

Some of the most striking data are those from the time-of-flight mass spectrometer designed at the Max-Planck-Institut für Kernphysik at Heidelberg, built with French funds and operated jointly with Soviet specialists. The instrument provides an instant mass spectrogram of the content of the plasma produced by the impact of individual dust particles on a silver target. Vega 1 provided more than 1,000 records, Vega 2 nearly half as many. Although many particles appear to have been composed largely of light elements with atomic number less than 20, there is a substantial fraction with components of atomic number between 30 and 50. The interest will be to see whether the second group shows the presence of silicon and iron, expected constituents of primordial dust.

The two encounters have also provided an extremely accurate determination of the orbit of Halley's comet, chiefly through the work of the Pathfinder project designed so as to provide the basis for guiding the European spacecraft Giotto to its encounter with Halley's comet on Thursday this week. On the basis of the data from Vega 1, the Pathfinder project says that it has been able to define the position of the comet in its orbit with an uncertainty of only 15 km.

One of the team responsible says that ground-based observations in the past few years have refined the orbital elements of the comet by a factor of 100, while triangulation based on accurate radio-interferometric measurements of the position of the Vega spacecraft and a knowledge of their pointing angles towards the Halley nucleus have made possible a further refinement by a factor of 30.

Elaborate arrangements have been made to communicate this information to those at the Darmstadt centre of the European Space Agency in West Germany, from where the Giotto encounter is being controlled. It is especially relevant that the uncertainty of the position of the comet along its orbit has been reduced from several hundred kilometres to a few tens of kilometres, given the ambition to send Giotto within 500 km of the nucleus.

A final decision about the direction of Giotto was due to be made on Tuesday this week. According to Dr Roger-Maurice Bonnet, the French scientific director of the Giotto project, the decision is certain to have been difficult. Five of the nine experiments on Giotto would prefer to get even closer to the nucleus than 500 km, but the other four wish to survive beyond the closest approach while two would like to survive by several days. One complicating factor is that nobody at

present understands why there has been such a rapid decrease of dust production in the three days between the two Vega encounters.

Dust appears to have played havoc with several of the Vega experiments. A low-frequency plasma detector, based on an antenna stretched between one of the booms projecting from the spacecraft and the solar panels, went out of action just before the first encounter, to the resigned annoyance of the group responsible. On the less successful encounter of Vega 2, several instruments went out of action before the encounter. One speculation is that the spacecraft may have become electrically charged in such a way as to affect the operation of instruments requiring a stable reference voltage.

The encounter of Vega 2 was also complicated by a pointing error, which for most of the approach had the cameras locked on to a bright patch of scattered sunlight offset from the nucleus of the comet. The most serious consequence was that the automatic microprocessor-based program relating the exposure of each frame to the brightness of the aiming point consistently over-exposed the nucleus. In a nail-biting emergency, the error was corrected only ten minutes before the encounter with the help of a back-up guidance system.

The general opinion of experimenters and managers is that the project has been outstandingly successful. Those gathered at the Institute of Space Science of the Soviet Academy of Science, which has mounted the exercise, have been quick to pay tribute to the international collaboration on which the project has been based.

Collaborators come from France and West Germany in Western Europe and from Czechoslovakia, Hungary and Yugoslavia in Eastern Europe. Although the United States is not formally involved in a project mounted when relations between the Soviet Union and the United States were in the doldrums, Professor J.A. Simpson from the University of Chicago has flown a novel dust detector on the identical Vega spacecraft at the invitation of Professor R.Z. Sagdeev, the director of the Soviet project and the institute at which it is based.

By common consent, the Deep Space Tracking Network of the US National Aeronautics and Space Administration (NASA) had also provided invaluable assistance both during the preliminary encounters of the Vega spacecraft with the gravitational field of Venus at the end of last year and in refining the trajectory of the vehicles.

Sagdeev recalls the Jet Propulsion Laboratory in California calling up one night with the question why one of the two Vegas had apparently vanished from the radio sky. □

Halley's comet

Telling the truth in real time

Moscow

As this monumental city last Wednesday watched its four-month covering of snow turn to grey slush, the builders of the first of four probes to Halley's comet were celebrating (mostly in the abstemious Gorbachev style) their successful first look at that eroding dirty snowball in the sky. But the more striking achievements may be the informality of the international collaboration on which the Vega project was based and the frankness with which all the details of the project were displayed.

To say that the outcome confirms the dirty snowball picture of a comet, that of a lump of frozen water with embedded dust, true though it might be, would be to diminish the occasion. Getting within 8,000 km of an object nine light-minutes away may be child's play these days, while Voyager has shown that instruments can be made to function at least a decade after being built, but who can suppress childish pleasure that five years of planning for an hour or so of proximity to Halley should, in the end, succeed?

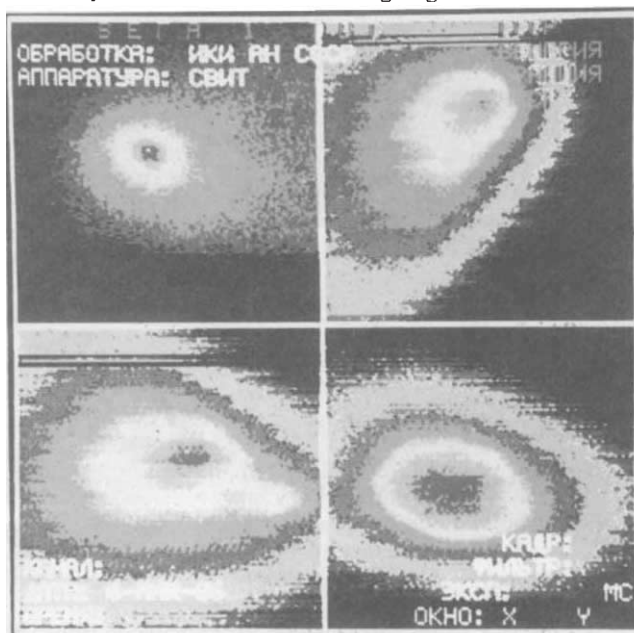
There were, in fact, two Vegas and thus two close encounters, on Wednesday last week and on Sunday. As launched (in December 1984), the two spacecraft were identical. Each of them dropped a balloon into the atmosphere of Venus before being deflected towards the retrograde orbit of Halley, still last week between the orbits of the Earth and Venus (but with the Earth on the other side of the Sun). At the Soviet Academy of Science's Institute of Space Sciences, the interval was spanned by a kind of perpetual polyglot symposium interrupted only by visits to other laboratories and the need for lunch.

Strictly, there have been two on-going

symposia. The experimentalists have been crammed with their microprocessors into a ground-floor room about 30 m² dominated by a full-scale model of Vega. Each experiment has been able to monitor its own data stream in real time; most, almost as a matter of pride, have thought it seemly to produce an instant analysis and interpretation as well. The Hungarian image processors were back-to-back with the Austrian recorders of the magnetic field. The Czechoslovak and US dust particle experiments were within whispering distance of each other (with the West Germans across the corridor). The French team taking ultraviolet spectra have been almost inaccessible in their corner.

The other symposium, on the floor above, has been for dignitaries, and has been recognizable by the bottles of mineral water on the tables. (Local opinion holds that "something is going to happen" when the bottles are changed; "they know we'd drink them if they left them there too long".) The dignitaries have included Dr Fred Whipple, lately retired from the Harvard-Smithsonian Astrophysical Observatory, the originator of the dirty snowball picture, Dr Thomas Donahue, chairman of the US National Academy's Space Sciences Board, delegations from the European Space Agency (ESA), the US National Aeronautics and Space Administration (NASA) and from the Space Sciences Institute at the University of Tokyo and vice-presidents of the Soviet Academy, who dropped in from time to time.

During both encounters with Halley's comet, this symposium has been kept thoroughly up to date by a display on two substantial video screens of data as re-



False-colour images of Halley's comet (shown here in black and white) from Vega 1 at around the time of closest approach at 07.20 GMT on 6 March, with four different filters. Although the top left-hand image shows a small peak of intensity (which appears white) near the centre, this cannot be identified with the nucleus of the comet, but is rather an artefact of the colour-coding system. The double structure of the image immediately below, in the near infrared, may signify the jet of dust encountered by Vega 1. The distance across each frame is about 1,500 km.

ceived and analysed by the experimenters in virtually real time. Professor Albert Galeev provided a running commentary, doing his best to interpret data as they appeared and succeeding reasonably well except when the projectionists switched displays in the middle of a sentence. (A Hungarian image processor says he had not been asked to provide this frill until a month ago.) It is an interesting and moving experience to see mass spectrograms of the products of the impact of single cometary dust particles on a silver target literally before the West Germans who built this ingenious apparatus had a chance to censor them.

Formally, the six-day symposium was a meeting of the inter-agency coordinating group bringing together the US and European space agencies, an amalgam of the two agencies in Japan and that from the Soviet Union. Academician R.Z. Sagdeev, the driving force behind Vega and the manner of its public presentation, wanted to use the occasion for hammering out a further programme of collaboration, offering to form a working group to say what should be done next before everybody took flight for Darmstadt and the Giotto encounter. But neither NASA nor ESA was ready for this proposal, suggesting instead a formal discussion at a meeting planned for Padua in November.

Apparently there is a plan to move from Padua to Rome for a public presentation of the Halley data, including an audience with the Pope, to whom each agency is expected to provide its four best pictures of the comet, possibly on 7 November. Those making the arrangements appear to have overlooked the significance of the date, the anniversary of the October Revolution. Sagdeev wondered whether the Pope might be persuaded to celebrate mass for the occasion.

For the rest, the dignitaries behaved as dignitaries do on these occasions. Their congratulations have been full-throated, not merely for the success of the two Vegas but for the openness of the proceedings of which they have been a part. The institute's staff has been surprised to find the proceedings televised live from the "data analysis room", which is said to be unprecedented. But what about the Party Congress? "That's the other exception."

The good-humoured meeting has also been sadly depressed about the shuttle tragedy, wondering what the effect will be on the US space programme. NASA officials say that it has been decided to plan on the assumption that the shuttle will be operating again after a delay of twelve months, and that priority will then be given to launching the Galileo mission to Jupiter and the Hubble space telescope. The snag, it seems, is that "it's costing us \$10 million a month to keep things on hold".

John Maddox

Halley's comet

Japan's probe rocked by dust

Tokyo

"It's great" was how researchers at Japan's Institute of Space and Astronomical Science (ISAS) reacted to news that their space probe *Suisei* (comet) had safely passed its point of closest approach to Halley's comet. *Suisei* was twice shaken by collision with high-velocity dust particles as it came within 150,000 km of the

some rapid calculations as to just how big the particles must have been to rock the 135-kg probe, bearing in mind that their relative velocity is something like 73 km per second. Initial estimates were that the dust weighed about 10 μg ; but recalculation came up with a figure of 1–2 mg, giving the particles diameters of 1–2 mm. If this proves correct it is a very big sur-



Sun side of Halley at 22.06 on Saturday, 8 March (Japan time) but in neither case was it thrown sufficiently out of orientation to affect data collection.

News of the collisions and data from the encounter did not reach Japan until Sunday, for the probe was on the wrong side of the Earth to communicate with ISAS's new deep-space observation centre at Usuda in Nagano Prefecture, west of Tokyo. Last commands sent to *Suisei* on Saturday were for the probe to continue to make observations with its ultraviolet camera and its ion-energy analyser according to programme. From Saturday evening until Sunday morning, signals were picked up at the US National Aeronautics and Space Administration (NASA)'s space observation centres in Spain and California, suggesting that all was going well. Then when *Suisei* came back round to the Japan side of the Earth, all data collected during the encounter were taken out of memory and transmitted to Japan.

The first and biggest surprise was that the probe had been shaken by collisions with dust particles just before the point of closest approach. The first collision at 20.54 knocked the rotation axis of the probe relative to the Sun out by 0.48°; the second at 21.26 added another 0.24° deviation. But neither was sufficient to affect observations of the comet. The first impact was also powerful enough to slow down the satellite's spin period of 9.184 seconds by 0.027 seconds; the second, presumably to a different part of the structure, had a scarcely noticeable effect.

News of the dust collision triggered

prise and may be bad news for the European spacecraft, Giotto, which is going very much closer to Halley later this week. "We don't know whether to call it good luck or bad", said one ISAS scientist of the collisions — bad luck to have been hit or good luck to have discovered something quite unexpected; no dust particles this large were expected to be thrown off by the comet.

Not unsurprisingly, there has been no time yet to analyse data in any detail. But results from the Lyman-alpha ultraviolet imaging camera confirms the variability in the hydrogen coma of Halley's comet that the Japanese team report in this issue of *Nature* (see p. 140). Detailed study should make it possible to measure the period of rotation of the comet. Preliminary results from the ion-spectrum analyser also indicate that a shock-like region forms on the Sun side of the comet as the solar wind is displaced around it. Ions from the comet were detected as far as 400,000 km away.

Suisei will continue to make observations of Halley for as long as possible. Over the next few days, *Sakigake*, the test probe launched in advance of *Suisei*, will come into position downwind of Halley to complement its observations. On board are instruments to measure plasma waves, ion temperature, velocity and density and the interplanetary magnetic field. But scientists at ISAS are already beginning to wonder what can be done next with *Suisei*. One of ISAS's satellites continued to function for six years and it is hoped that further useful observations can be made with the probe as it continues to circle the Sun.

Alun Anderson

Biotechnology

All systems go in Britain

It has been a fortnight of activity at the business end of biotechnology in Britain. First Genentech, the Californian company, was granted a British patent for its tissue plasminogen activator (TPA), which is expected to find a place in the treatment of myocardial infarction. Then the British pharmaceutical company Wellcome plc, while awaiting Genentech's announced intention to file suit against Wellcome for infringement of the newly granted TPA patent, set in motion an attempt to revoke the patent. And finally, the first British licences for a clinical use of interferon were granted.

The interferon licences are less of a coup than they might seem, since they are only for the treatment of hairy-cell leukaemia, of which there are fewer than 100 new cases a year in Britain. Interestingly, the Committee on Safety of Medicines, which grants the licences, has chosen not to discriminate between Wellcome's interferon (trade name Wellferon), which is purified from a cultured cancer cell line and is a mixture of the many slightly different α -interferons, and Intron A, which is a single α -interferon produced in bacteria by recombinant DNA technology by Schering-Plough (the licence goes to the company's British subsidiary Kirby-Warrick Pharmaceuticals Limited). Hoffmann-La Roche has

also applied for product licences for its recombinant DNA-produced α -interferon (Roferon A).

It remains to be seen, said Wellcome officials last week, whether there will be conditions in which their broad-spectrum product is advantageous or disadvantageous compared with single interferons. So far, both Schering and Hoffmann-La Roche have each concentrated on only one α -interferon. Dr Daniel Catovsky of the Royal Postgraduate School of Medicine in London, who has carried out a trial of Wellcome's interferon, said this week that it seems as if all cases of hairy-cell leukaemia seem to benefit in a predictable way from treatment.

Schering's interferon- α_{2b} started life as a clone in Dr Charles Weissmann's laboratory at the University of Zurich and was developed by Biogen for Schering, which now has a variety of product licences in the Philippines, Belgium and Ireland but none yet in the United States. Hoffmann-La Roche's interferon- α_{2a} was originally cloned by Genentech.

In contrast with its policy on interferon, Genentech has chosen to be much more independent with TPA, although it does have marketing agreements with some overseas companies. From the first cloning of TPA's DNA at Genentech in 1982 (starting with a TPA-producing cell line characterized by Dr Désiré Collen, see below), the company had reached the stage last year of converting a car park into a fermentation plant dedicated to the culture of the mammalian cells that produce TPA after its DNA has been genetically engineered into them; several European patents have already been granted and clinical trials are far advanced.

By contrast, Wellcome's TPA is still at preclinical stage and is far from being a home-bred product. The original cloning was carried out by Joe Sambrook at Cold Spring Harbor Laboratory in conjunction with the Cell Biology Corporation which was a collaboration between the laboratory and Baxter Travenol. Pre-production development was carried out for Travenol by the biotechnology company Genetics Institute, based in Cambridge, Massachusetts. And Wellcome then took over for the production stages, because of its expertise in the large-scale culture of mammalian cells gained during its interferon work.

The hope is that TPA will be a better thrombolytic agent, particularly in the removal of clots in the coronary arteries, than the currently favoured streptokinase. Commenting on the clinical trials in *New England Journal of Medicine* last October, Dr Sol Sherry of Temple University School of Medicine concluded that "se-

Plutonium dumping?

REPORTS of plutonium on the seabed near the Loviisa nuclear power station in Finland have been greatly exaggerated, according to Erkkela of the Finnish Centre for Radiation and Nuclear Safety. But the reports have come at an embarrassing moment for Finland's nuclear energy industry. The two nuclear power companies, Teollisuuden Voima (TVO), which has two nuclear power stations at Olkiluoto, and Imatram Voima (IVO), which operates the two stations at Loviisa, are at present negotiating a cooperation arrangement in order to increase their combined generating capacity by 1,000 MW.

According to Ikus, the Centre for Radiation and Nuclear Safety monitors the whole Finnish coastline, but concentrates, naturally, on the vicinity of Olkiluoto and Loviisa. The reports about Loviisa referred, he said, to a plutonium concentration of about 100–300 becquerels per m², which, he said, is "typical for all our coastal stations". Helsinki radio, which first carried the story, had, he said, simply mistaken its newsworthiness.

In any case, Ikus said, the plutonium in question came, not from Loviisa, but from the fallout from nuclear tests in the past. When asked about reports reaching Sweden from Estonian dissidents that nuclear waste from the Paltiski submarine base is being dumped in the forest near Tallinn or discharged directly into the Baltic, Ikus said that he was "sure on scientific grounds" that this was not the case with the plutonium found in the Finnish sediments.

Vera Rich

Swiss prize shared

DÉSIRÉ Collen of the University of Louvain, Luc Montagnier of the Institut Pasteur and Michael Berridge of the University of Cambridge have shared this year's £600,000 prize awarded by the Louis Jeanet Foundation for Medicine.

The prize, named after a resident of Geneva who died in 1981, is awarded annually to up to three "outstanding and productive research workers" to provide them "with substantial funds to pursue their researches".

Collen, whose work on the mechanisms of thrombus dispersal forms the basis of the British patent for tissue plasminogen activator, awarded to Genentech two weeks ago, will use the prize money to apply genetic engineering to the production of new clot inhibiting or dissolving substances. The award to Montagnier, who discovered the virus that causes acquired immune deficiency syndrome (AIDS), is to be used to continue his research into this and other human retroviruses. And Berridge will continue to explore the presence and function in cells of substances related to inositol trisphosphate, the work for which his share of the prize was awarded.

Peter Newmark

rious questions remain about how much of a therapeutic advance it will represent as compared with the agents already available". Undeterred, Genentech and Wellcome are prepared to foot the bills for what might be a protracted test case for the patentability of the products of biotechnology. "If litigation does proceed", says Stephen Crespi, controller of patents for the British Technology Group, "it will shed some much-needed light in patentability in this very untested field."

Peter Newmark



Dr James C. Fletcher, administrator of the US National Aeronautics and Space Administration between 1972 and 1977, who has been invited to resume the post in succession to Mr James Beggs, who has resigned. The appointment is subject to approval by the US Senate. □

Prizes

US winners for Israeli prizes

Rehovot

US CITIZENS walked off with most of this year's Wolf Prizes, just as they did with most of the Nobel Prizes (although the prize categories are not identical).

In mathematics, for example, the \$100,000 Israel-based international award went to a pair of European-born Americans, Professor Samuel Eilenberg of Columbia University and Professor Alte Selberg of the Institute for Advanced Study at Princeton. Eilenberg was chosen for his fundamental work in algebraic topology and homological algebra, while Selberg was honoured for his profound and original work on number theory and on discrete groups and automorphic forms.

Opening up a new field of research, the study of chaos won Wolf Prizes in Physics for an American, Professor Mitchell Jay Feigenbaum of Cornell University, and a Frenchman working in the United States, Professor Albert Joseph Libchaber of the University of Chicago.

Professor Elias James Corey of Harvard University and Professor Albert Eschenmoser of the Swiss Federal Institute of Technology (ETH) share the Wolf Prize in Chemistry, in recognition of their research on the synthesis, stereochemistry and reaction mechanisms for the formation of complex natural products, particularly vitamin B₁₂. Eschenmoser is the first Swiss national to receive a Wolf Prize.

In agriculture, the prize is shared by Dr Ernest R.R. Sears of the University of Missouri and Sir Ralph Riley of the UK Agricultural and Food Research Council. The two cytogeneticists were cited for their independent research in basic plant breeding, which has permitted the insertion into wheat of genes representing desirable characteristics from alien plants so as to improve wheat germ plasmas and cultivars.

The Wolf Prize in Medicine, the last to be announced, went this year to a Japanese researcher, Professor Oshmu Hayaishi, president of the Osaka Medical College, for his discovery of oxygenases, a group of enzymes that are catalysts in the use of oxygen by living organisms. They also, as the citation noted, play a crucial role in the metabolic disposal of foreign compounds such as drugs and insecticides.

The Wolf Foundation, which has been awarding annual prizes since 1978, was established by the late Dr Ricardo Wolf, a German-born chemist and philanthropist who lived for many years in Cuba. Appointed Cuban Ambassador to Israel in the 1960s by his friend Fidel Castro, he remained in Israel after Cuba broke off diplomatic relations in 1967 until his death in 1981.

Nechemia Meyers

US science and technology

McTeague looks to the future

Washington

CONTEMPLATING the future of your agency when it is faced with a 28 per cent budget cut is not a cheerful prospect. But that is the plight of John McTeague in his first weeks as acting director of the White House Office of Science and Technology Policy (OSTP). "We can operate with a budget at almost any level, the question is which activities the office will undertake", says McTeague. "That is a matter of consideration by the Congress."

Last month McTeague was on Capitol Hill before a Senate appropriations subcommittee. While he did not appeal for more money for his agency, McTeague did say that he would be very happy with a 10 per cent cut in his budget, considering the alternative he is faced with.

McTeague is not behaving like a temporary incumbent marking time while waiting to be replaced. Before the congressional committee he was quick to point out the need for a stronger hand in guiding science and technology policy by the federal government. He plans to provide some of that guidance through his own role as chairman of the dozen committees making up the Federal Coordinating Council on Science, Engineering and Technology (FCCSET). While he refuses to state explicitly that he favours a cabinet level department, he clearly feels the present system is awkward.

McTeague considers that most people fail to appreciate the scope of OSTP's activities. "OSTP is a little like an impressionist painting", said McTeague. "If you don't stand back, you miss the picture."

One FCCSET committee, formed last December, is attempting to place renewed emphasis on international cooperative agreements in which the United States is involved.

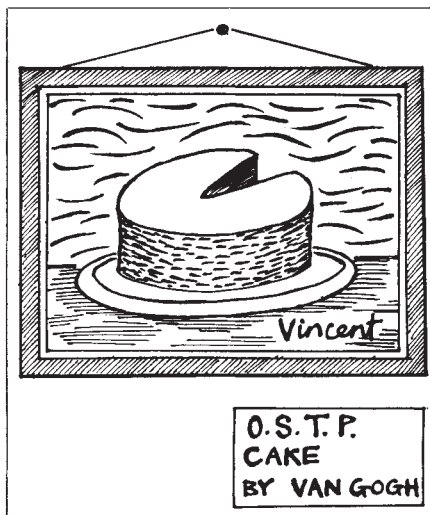
"It's not an effort to dictate individual programmes", says McTeague. But this does not mean that OSTP will support US participation in international programmes that do not fall in line with domestic priorities or international policies. McTeague believes that the most effective international interactions are those that take place informally, with government maintaining a hands-off policy.

One project with enormous international implications is the proposed Superconducting Supercollider (SSC). With a price tag approaching \$6,000 million, there is a natural reluctance to forge ahead, although McTeague maintains that the present administration firmly supports high-energy physics.

"We have to take a good hard look at it [SSC] and see how it fits in at this time with national priorities and budget exigencies", says McTeague. Although

greater cooperation with the European Organisation for Nuclear Research (CERN) may be a practical alternative to building SSC, it seems to fly in the face of a strong commitment by the government to maintaining "this nation's forefront role in high-energy physics".

OSTP has also recently become embroiled in the question of how to proceed with the US space programme following the shuttle disaster. The administration is faced with an expensive problem of whether to buy a new shuttle orbiter in a year when budgets are austere all round. While the administration has not yet officially made a decision, McTeague says expanded use of expendable rocket launch vehicles will provide a solution only "in the short run".



McTeague feels that the most effective federal role in science and technology "is not to attempt to drive specific commercial technologies, but to support forefront basic research, which will have payoffs not foreseeable at the time of funding". At the same time, he supports "problem focused" science centres at universities, aimed at areas of "broad national needs and relevant to industrial technology". McTeague explains this apparent contradiction by pointing out that it will be scientists, not the government, who will identify these "problems". The government's role will be to nurture these centres once their direction has been plotted.

McTeague has had three years in Washington to watch and participate in the workings of OSTP. He clearly intends to continue pressing for federal support of science, although he is anxious that industry should contribute its share to the financial burden. With the Gramm-Rudman cloud hanging over the United States and the federal budget, running OSTP will not be an easy job. Perhaps it never is.

Joseph Palca

Strategic Defense Initiative

Grilling for Abrahamson

Washington

A RARE glimpse of progress and problems with President Reagan's Strategic Defense Initiative (SDI) was offered last week by Lt General James A. Abrahamson, chief of the SDI organization, baited by a feisty Senator William Proxmire (Democrat, Wisconsin), who kicked off the congressional SDI appropriation hearing by declaring "if you expect a 75 per cent budget increase you're in for a rude awakening".

Abrahamson exercised commendable restraint as he outlined changes to the programme since his last appearance before the defence subcommittee of the Senate Appropriation Committee. Space-based chemical lasers, once one of the SDI organization's high hopes, are out; spaced-based kinetic kill vehicles are in. And point missile defence, a more restricted goal than President Reagan's concept of a shield covering the whole United States, will "need study".

The administration's proposed budget for SDI in fiscal year 1987, which starts next October, is \$4,800 million: the current budget is \$2,760 million, 25 per cent less than the President had requested. The rough treatment meted out to Abrahamson by some members of the subcommittee

last week owes something to the fact that, alone among military programmes, SDI has been totally exempted from cut-backs in the current financial year resulting from the Gramm-Rudman deficit reduction law.

But Proxmire — who dominated the proceedings though not formally in charge — also ridiculed the SDI organization's decision to classify a report on SDI contracts prepared by the General Accounting Office entirely from unclassified sources.

Declaring that all the information in the now classified report was readily available through other official and private publications, Proxmire demanded whether Abrahamson thought "the Soviets haven't got \$625 (the cost of the documents) and a calculator". Abrahamson thought it was not in the national interest to have details of SDI contracts and contractors listed in a single, easily obtained volume.

Abrahamson went on to unburden himself with some admissions of greater than expected technical problems that have been encountered in discriminating real warheads from decoys. The SDI organization's "red team" (set up to think of ways in which the Soviet Union might counter SDI) has made US versions of Soviet warheads and possible decoys so successfully that the blue team was forced to conclude that "passive discrimination is not quite sufficient" to tell them apart. Undeterred, Abrahamson explained how he now hopes to use low-power neutral particle beams and "cheap" detectors for "interactive discrimination" — although "more research is needed at this time".

Research on space-based lasers has been slowed "with a profound sense of technical regret" because of budget pressures, Abrahamson explained, in order to avoid accepting slippage in the more promising parts of the programme.

Proxmire was unforgiving. After reminding Abrahamson that the organization had cancelled a \$62 million contract for space-based infrared sensors because costs had exceeded estimates (and settled for a less capable sensor), he demanded to know "why Congress should believe the priorities won't change again" and whether it was not that the SDI organization had come to realize that space-based weapons were militarily vulnerable. Abrahamson replied with an allegory: just because aircraft are vulnerable to attack they have not become militarily useless, he said. Anyway, satellites could be defended but details were classified.

Abrahamson averred that in other areas research had been more successful than expected. (Very fast rockets in pods in space apparently look good.) But he did

provide one hint of where the SDI organization might be going if indeed, as many critics contend, an overall missile defence shield over the United States proves technically too daunting, although this was not a possibility Abrahamson thought at all likely.

While a decision on the feasibility of the population shield should be possible by the mid-1990s, according to Abrahamson, a decision on a more restricted goal of point missile defence should be possible somewhat earlier, perhaps as soon as the early 1990s.

Tim Beardsley

Canadian budget

New twist

CANADA has introduced a "brand new twist" to the budget process in an attempt to increase money for research, and provide university research councils with "stability of funding for the next five years". But many, not least the research councils themselves, are not too clear on how this stability is to be achieved.

The new scheme is designed to increase private sector support for research. Starting in fiscal year 1987, the government will match dollar for dollar contributions from private industry to research councils, up to a total of 5 per cent of an individual council's budget. The problem is that nobody is sure how the matching scheme will operate in practice, and the government is still working on an explanation.

Moreover, after increases for Canada's three research councils in the first year of the five-year plan, the government's core contribution drops, and any increase will have to come from the matching funds scheme. Uncertainties about how much this will amount to has forced the Natural Sciences and Engineering Research Council (NSERC) to cancel plans for expansion, and even at current levels NSERC will have to cut back its programme.

If the private sector does embrace the plan, the budgets of the research councils will increase by \$1,000 million over the next five years. The government is in the process of overhauling the tax structure to encourage private investment in research council projects.

A joint statement issued by the Canadian Association of University Teachers and the Association of Universities and Colleges of Canada expressed "grave concern" about the government's plan. While not opposed to closer cooperation with business, the university groups worry about how the relationship with industry will be defined.

A big loser in the budget battle is the Science Council of Canada. The council's budget was cut by half to \$2.5 million, and its full time staff was reduced from 68 to thirty.

Joseph Palca

SSC in doubt

Washington

OBSERVERS on Capitol Hill are becoming suspicious that the US Department of Energy's possible mammoth accelerator project, the 60-mile-around Superconducting Supercollider (SSC), may be a hostage to the Gramm-Rudman deficit reduction act and perhaps to the cost of replacing the space shuttle *Challenger*. Recent testimony given by representatives of the department makes clear that no decision has yet been taken even to support continuing research in fiscal year 1987 (which starts next October).

The department insists that a decision on whether to build SSC, estimated to cost \$6,000 million, will be made only after it reviews a comprehensive technical report due from the Central Design Group on 1 April.

But the tone is decidedly less enthusiastic than in previous years. Project managers had hoped for \$65 million to conduct research on SSC next year, but even if all goes well the total is likely to be rather less than \$40 million, to be obtained by reprogramming from elsewhere in the high-energy physics budget. In any event, the project is big enough to require cabinet level approval before it goes ahead. Actual construction funds could not be expected before fiscal year 1988.

Tim Beardsley

Nuclear winter

Has winter become fall?

Blacksburg, Virginia

Two years of intensive computer modelling have substantially refined the concept of "nuclear winter" but the diverse opinions on its strategic and diplomatic implications suggest that any real impact on US defence policy is still some years away. A conference* last week that included representatives from universities, government and the military heard sharply differing views on the importance that should be attached to the results so far obtained.

There was a consensus that although models have become increasingly sophisticated and useful, there is still great uncertainty about the source terms for the quantity of dust and smoke that would be injected into the atmosphere following a nuclear exchange. Beyond that, pre-established positions dominated. Colonel Terry Hawkins, representing the Pentagon, agreed that the possibility of nuclear winter was a "sideshow" that was simply another reason to avoid nuclear war, and repeated the familiar Pentagon line that the present policy of deterrence through strength is the best interim policy. But Hawkins did use the idea of nuclear winter to support a transition to defensive weapons as a long-term aim through the Strategic Defense Initiative (SDI). Others, believing SDI to be destabilizing, argued exactly the opposite.

Robert Simmons of the Arms Control and Disarmament Agency agreed that the threat of nuclear winter was a further reason for disarmament. Maurice Roesch III, assistant director of the White House Office of Science and Technology Policy (OSTP), said that it was necessary to understand the phenomenon "before we go too far trying to understand policy implications" and made clear his view that understanding had not yet been reached.

Others, such as Thomas Malone, former secretary-general of the International Committee of Scientific Unions' Scientific Committee on Problems of the Environment (SCOPE), emphasized that nuclear winter had attracted much interest in non-combatant countries such as India, which would in future take a much greater interest in nuclear war doctrines of the superpowers and catalyse "worldwide moral indignation" over nuclear weapons.

SCOPE last year produced the most thorough study yet of climatic effects of nuclear weapons. But George Rathjens of Massachusetts Institute of Technology strongly criticized "irresponsible" misrepresentation of the science of nuclear winter ("the worst example in my memory"), while making it clear that he was referring to popular press coverage and public statements by the populist scientists rather than to the research papers themselves.

And some pointed out that a nuclear war between the superpowers would be quite damaging enough to developing countries because of the loss of trade, even without nuclear winter.

Ambassador Richard Butler, Australia's representative at the Geneva arms talks, made an impassioned plea for the superpowers to live up to their commitment to abandoning nuclear weapons altogether and said the changes now occurring in the Soviet bureaucracy make it possible to start work on "a new fabric of security for the forthcoming post-nuclear age".

More sophisticated recent three-dimensional interactive global circulation models tend to predict smaller temperature drops in continental interiors than the one-dimensional model of Turco, Toon, Ackerman, Pollack and Sagan (*Science* 222, 1283-1292; 1983).

The latest results from the global circulation model at the National Center for Atmospheric Research at Boulder, Colorado, continue the trend. Preliminary results reported by Stephen Schneider and Starley Thompson at a Defense Nuclear Agency technical review from a model incorporating microphysics effects to coagulate smoke and dust particles, remove them by precipitation and allow the changed particle size distribution to affect radiative transfer terms, predicted an average temperature drop in the Northern Hemisphere for a 180 terragram smoke injection at close to 12 degrees centigrade, about 3 degrees less than a similar model without the microphysics package. The results are likely to be referenced in a forthcoming but delayed report on nuclear winter requested by Congress from the Department of Defense.

Russell Seitz, currently at Harvard University, quoted Schneider as referring to the new simulations (together with new estimates of fuel loading by George Bing of Lawrence Livermore Laboratory that suggest 60 terragrams would be a more plausible smoke injection for a medium-sized conflict) as indicating that "nuclear fall" would be a more appropriate description of climatic changes after a nuclear exchange.

Richard Small of Pacific Sierra Research Corporation presented an analysis of fuel loadings resulting from a strictly counterforce exchange that predicted a maximum of 3 terragrams of smoke in the atmosphere, not enough to produce any nuclear winter effect. But Small agreed it was "prudent" to consider policy implications.

George Carrier of Harvard University said that trends in modelling suggest that for a nuclear exchange in the Northern

Hemisphere to devastate agriculture in the Southern Hemisphere "almost all the as yet unknown parameters will have to come out on the serious side", unless agricultural systems are even more sensitive to climate effects than is now recognized.

Others drew attention to other possible long-term consequences of a nuclear exchange. Michael McCracken of Lawrence Livermore Laboratory said that precipitation over land might be significantly reduced by even relatively small smoke injections, possibly leading to a failure of the monsoon.

Mark Harwell of Cornell University, a co-author of last year's SCOPE report on nuclear winter, reiterated that even the more modest temperature drops predicted by the latest models would still have disastrous consequences for agriculture given the effects of spatio-temporal variation, and pleaded for more research support for biological effects. The great majority of US federal support for nuclear winter research (nominally \$5.5 million in the current fiscal year) is for geophysics work, although the National Science Foundation still has \$100,000 uncommitted that could be diverted to biology.

Much discussion centred on the proposal, put forward by Senators Proxmire and Hatfield, that the United States and the Soviet Union should establish a joint commission to study nuclear winter. Alan Hecht of the National Oceanic and Atmospheric Administration, who drew up the US research plan before the topic was taken over by OSTP, spoke favourably of the potential of Soviet scientists to contribute to nuclear winter research, although the disappearance a year ago of Soviet researcher Vladimir Aleksandrov had done nothing to make collaboration easier. But he complained of the difficulty of gaining access to top Soviet scientists.

Large-scale experiments on fires are thought to be one area where collaboration might be fruitful; observations of deliberate forest burns are also being planned with the Canadian Forestry Service. Work on computer modelling is complicated by the Department of Defense's insistence that Soviet citizens should not have access to US supercomputing expertise, currently a source of friction between the Pentagon and the National Science Foundation. Administration officials strongly prefer to work on developing lower level contacts than a commission would imply, however, most probably under the auspices of the 10-year old US/Soviet bilateral agreement on environmental cooperation, which has a working group on atmospheric research chaired by Hecht. A final decision on whether to pursue that option has yet to be made.

Tim Beardsley

* Nuclear Winter: strategic implications, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 6 March 1986.

South Africa and science for all

SIR—Certain critics of South Africa are now threatening to invade the arena of the natural sciences. The objects of this vociferous minority, in essence a congerie of ill-informed scientists and non-scientists, are to isolate our scientific community by impairing the free flow of scientists and scientific information to and from South Africa. The latest manifestation of this campaign is the attempt to preclude the attendance of South African scientists at the World Archaeological Congress.

I am the president of the South African Council for Natural Scientists, an autonomous body created by statute, which consists of eminent scientists nominated by their peers from all the branches of the natural sciences.

The council's terms of reference are to register adequately qualified natural scientists and to promote the interests of their profession; to protect public health, safety and interests generally against actions by inadequately qualified or non-qualified persons who venture into the natural scientist's field; and to administer a code of professional conduct for all registered natural scientists.

The council, which is totally non-racial, issued a policy statement early in 1985 in which it reaffirmed its attitude in support of the universality of science. It further expressed itself in favour of the free and unfettered pursuit of science and reasserted its constant aim to promote the interests of all natural scientists, irrespective of race, colour, creed or sex.

The proponents of South Africa's scientific isolation find justification (*sic*) for their actions in convoluted political arguments which are spurious in the very context of what science is all about: the pursuit of knowledge and the advancement of human capabilities. Surely scientific contribution or discovery cannot in the least be discounted or deemed irrelevant or less cogent merely because it emanates from a particular individual or country.

International boundaries and ideological barriers are totally irrelevant in the pursuit and exchange of knowledge. Scientific genius crops up without regard for country, language and race. The laws of nature are identical in all lands: they are there to be discovered and applied, accessible to the researchers of all nations. The natural scientist constantly endeavours to think in terms of realities that transcend parochial interests and personal aggrandizement: he is, in fact, a member of an international community. This we believe, and would commend to those attempting to disrupt the community of natural scientists.

It is common cause that scientists worldwide are, by and large, men and women who are intelligent, objective and fair-minded. In the particular context of the

subject matter of this appeal, these attributes have significantly come to the fore of late: large numbers of letters are being received by scientists in this country from their colleagues worldwide in which the pernicious activities of our would-be isolationists are severely slated.

V. PRETORIUS

*South African Council
for Natural Scientists,
Private Bag 11303, Brooklyn,
Pretoria 0011, South Africa*

Chinese forest

SIR—In the article "South China's germ-plasm tank" (*Nature* 318, 220; 1985) there is a reference to what must be the Ding-hushan Forest Nature Reserve, an out-station of the South China Institute. It is not, as described, a 1,200-hectare arboretum and if it were, it would be by far the largest arboretum (tree collection) in the world. It is a largely untouched relic of natural subtropical forest which has remained so because it surrounded a temple and was regarded as sacred ground. Nor is it 1,200 hectares, but 300 hectares surrounded by 900 hectares of recent *Pinus massoniana* plantation buffer zone.

Your article says that Dinghushan may become one of the UNESCO Man and the Biosphere projects. It has been one for some time and this is proudly proclaimed at the entrance to the reserve.

E.H.M. HARRIS

*Royal Forestry Society,
102 High Street,
Tring, Hertfordshire HP23 4AH, UK*

Bigfoot no joke

SIR—As the leader of the "Bigfoot conservation group" that threatened to disrupt Mark Keller's armed Bigfoot hunt in 1984 (see *Nature* 313, 418; 1985) as well as a Loch Ness monster investigator (1982 and 1983), I should like to comment on what you say. First, our group wanted to stop Keller because we had no faith that he would not shoot some unarmed hiker in error. Then, we (Project Bigfoot) found almost no commercial interest in Bigfoot by chambers of commerce in the north-west. The local citizens simply wanted to stop hikers and others from running any risk in their woods from trigger-happy Bigfoot hunters. Bigfoot himself and his tribe seem to be able easily to avoid armed men, as Keller found after four months of no results.

But your main complaint seems to be that too much money and public commission time is being spent on what the writer feels is wasted effort. He wants that money/time spent on serious science, such as anti-pollution studies. In response, let

me suggest that Bigfoot/Nessie/Chessie research may in time uncover some serious scientific information on the nature of energy-forms that we occasionally see but cannot catch or kill. The borderlands of science are often seen as a waste to study, but today's nonsense can often become tomorrow's serious scientific achievement.

ERIK BECKJORD

*The Cryptozoology Museum,
18711 Pacific Coast Highway,
Malibu, California 90265, USA*

Animal intelligence

SIR—If man can be separated from animals only because he shows foresight, what can we make of the squirrel burying nuts in advance of the cold weather? Or the ants that show true agriculture with their fungus gardens? Andrejs Baidins' rule of thumb (*Nature* 319, 172; 1986) is as shaky as any other.

R. POWERS

*Department of Palaeontology,
British Museum (Natural History),
Cromwell Road, London SW7 5BD, UK*

Acid rain from jets?

SIR—Any motorist knows how corrosive exhaust fumes are on metal. No doubt aeroplane exhaust gases have much the same property, but since much of the emission from modern civil aircraft occurs at high altitudes, the vapour trails of droplets of ice containing any corrosive material must travel hundreds of miles to reach ground level. Where they eventually arrive will be governed by the variable weather conditions that they will encounter in falling, or being washed down in rain.

It would be of considerable interest and practical importance to take samples of air containing water vapour and dissolved substances from the vapour trail area of the flight path of one of the large civil jets and to analyse the samples for "corrosive" content.

The present conventional hypothesis for the "acid rain effect" puts the blame on the water vapour contamination from the smoke stacks of industrial plant. But present coal consumption is only a little more than one-third of what it was in the 1900s, and even though oil firing is much increased since then, it seems unlikely that this source could be the main factor responsible for the quite recent rapid worsening of the mortality of trees due to acid rain. If this was the main cause, why were not its effects much more severe in the early years of this century? Atmospheric pollution on a large scale due to aircraft exhaust fumes is a much more recent phenomenon; this may be the cause.

FRANK PYGOTT

*Navigation Cottage, Willoughby,
Rugby, Warwickshire CV23 8AG, UK*

A child's guide to Soviet science

The openness with which the Soviet Vega project was conducted at the weekend is both a pointer to the nature of Soviet science and a sign of how fruitful collaboration might be extended.

Moscow

WHAT can be learned about Soviet science in a mere five days in Moscow? Not much, of course. But the way in which the Vega project has been conducted in the past week, with quite astonishing frankness and friendliness, points to some important questions about the nature of the enterprise, and suggests some cheerful answers. The legend, in the West, that Soviet science is locked from sight of others behind a veil of paranoid secrecy is plainly, at best, a half truth.

The Vega project is the creation of one man, 54-year-old Academician R.Z. Sagdeev, who has been director of the Soviet academy's Institute of Space Sciences since 1973. He is Russian in the energetic mould, small, wiry and with a wry tongue. Part of the way through an account last week of how his laboratory has been simulating the effects of dust particles on metal targets, he could not suppress a reference to "other people's" interest in accelerating particles to high velocities "for other purposes". Last November, Sagdeev was the only scientist among the party of Soviet advisers at the Geneva summit.

The origins of the Vega project are interesting in themselves. Several people claim part of the credit. The Vega launchings were originally scheduled chiefly as vehicles for a collaborative experiment with the French for the exploration of the atmosphere of Venus by means of balloons. In the event, it seems to have fallen to the then director of science at the European Space Agency (ESA), at a three-hour meeting at Moscow airport, to persuade Sagdeev "to fly to Halley as well".

The project has been thoroughly international, with no trace of chauvinism. Collaborators from elsewhere have been invited on the basis that their instruments have a valuable contribution to make. J.A. Simpson, from the University of Chicago, for example, says that his instrument (which records dust particles penetrating an electrically anisotropic dielectric between two metal foils by means of the nanosecond voltage pulses they generate) is on board because a Soviet scientist heard him describe the instrument at a meeting; the invitation to Vega followed quickly, and was irresistible.

Other collaborators speak warmly of the informality of the collaboration. Dr Jochen Kissel from Heidelberg, whose dust detector is a mass spectrometer of novel design, says that Sagdeev's institute

has been easier to deal with than the US Aeronautics and Space Administration (NASA) and even ESA, both of which require more paperwork. This is especially interesting because relations between the Soviet and West German governments do not officially exist (although there is an agreement on scientific collaboration between the Soviet academy and its West German counterpart, the Deutsche Forschungsgemeinschaft), which is why the collaboration has had to be laundered through France.

Kissel also says that the Soviets have been more daring than other agencies would have been in their willingness to fly instruments not previously used. This applies both to his present instrument and to that on which he is collaborating with Sagdeev's institute to produce a chemical analysis of the surface of the martian satellite Phobos by firing a powerful laser at the surface and collecting samples of the plasma thus generated, all this in the few minutes during which the instrument will be near the satellite, ideally a few metres up.

Not all Soviet institutes have this reputation for daring, of course. Part of the explanation at Sagdeev's institute is Sagdeev himself; his colleagues say that he has made a tremendous difference to the spirit of the place, especially by the recruitment of able scientists from elsewhere. (Even so, people say, there remains some of the dead wood that is the bane of the Soviet research system.) Another is the literal nature of the Soviet regard for the material world, of which science is a description; if there are good theoretical reasons for supposing that some phenomenon should occur, what reason can there be for behaving as if it will not? The Soviet educational system may also help; the way that the better university students spend two or three years at an institute before their first degree, often in the expectation of staying on, while frequently the reason for the accumulation of dead wood, can be in the hands of a daring director a way of finding firebrands.

Sagdeev's successful pursuit of international collaboration seems also to be an illustration of the independence that the director of a Soviet institute can enjoy. I have no means of knowing what the chauvinistic pressures may be, although the Soviet space programme has traditionally made a point of involving others through its "Interkosmos" programme.

So too has the US programme in its time. But the West is now mostly in the cost-sharing mode; collaborators pay their share of the overheads as well as their direct costs. Sagdeev has obviously been able to make the most of Soviet readiness to be flexible on these points. Other institute directors could no doubt follow suit if they were so inclined. It will be interesting to see if Sagdeev's example rubs off.

The other striking feature of the past week has been the frankness with which the details of the Vega project have been displayed. Visiting dignitaries were at first somewhat startled to find themselves expected to discuss where collaboration should go next (solar-terrestrial physics seems to be the favourite) in the presence of journalists. Most experimenters took well to the notion that they should provide virtually a kind of running commentary on their data.

The proceedings have been starkly in contrast with the way that even NASA handles these affairs. (The NASA official who said in public that "even we have something to learn" knew that he had blundered before the word "even" had escaped his lips.) But this is not the first time that the Soviets have startled the West by frankness. In 1958, for example, V.A. Kurchatov astounded a British audience at the Atomic Energy Research Establishment at Harwell with a full account of the Soviet thermonuclear programme, from which stemmed the still-continuing international collaboration in that field. Kurchatov was succeeded as director of the Soviet establishment at Dubna by the even more outspoken Artsimovich, himself a one-time colleague of Sagdeev.

It is far from fanciful to think this lineal connection between people who know the value of frankness is significant. Moreover, it is not necessary to suppose that there has been a high-level decision by the Soviet government that the scientific part of the space programme should be made "open". As the French have known for some time, it has always been more accessible than it may have seemed. On this occasion, Sagdeev's personal interest, and his success, seem to have been the main-springs of what appears to have been a change (but which may be simply a mark of a neglected opportunity). This, obviously, is another example that other Soviet directors, and potential partners in the West, may follow. **John Maddox**

Lasers

Fusion turned inside out

from R.G. Evans and M.H. Key

A GROUP of researchers from Osaka University and a compatriot at Bell Laboratories have recently proposed a rather novel approach to controlled fusion (Hasegawa, A. *et al. Phys. Rev. Lett.* **56**, 139; 1986). They have sought to combine the advantages (or perhaps the pessimists would say, the disadvantages) of both laser-driven inertial confinement and magnetic confinement fusion, and their approach makes use of the highly efficient CO₂ laser. This type of laser has been increasingly dismissed as a fusion driver because of its long wavelength, as we recently discussed in these columns (*Nature News and Views* **313**, 94; 1985).

Impressed by the ease with which lasers produce megagauss magnetic fields when they interact with solid targets (for an explanation, see Key, M.H. *Nature News and Views* **276**, 210; 1978), the authors seek to harness these fields in much the same way as a conventional magnetic confinement device such as a Tokamak. The magnetic field *B* in a Tokamak or mirror

Alamos National Laboratory in the United States, has been devoted to devising a fusion scheme that makes use of CO₂ lasers, which have the necessary efficiency and energy capability for a fusion driver, but no satisfactory solution has yet been proposed. The present novel target design perhaps offers such a solution.

The figure illustrates the principle in which a CO₂ laser is fired into a hollow target. The copious energetic electrons produced by the CO₂ laser form current loops (as shown) and produce magnetic fields of several megagauss which spread along the walls as electrons are driven outward by a force caused by perpendicular electric fields *E* and *B* (Forslund, D.W. & Brackbill, J.U. *Phys. Rev. Lett.* **48**, 1614; 1985). Both laser radiation and electron heat conduction inside the cavity ablate material from the inside walls and the relatively cold material flows towards the centre. Here the temperature is higher because of the laser heating and it is suggested that fusion temperatures (10–100 million degrees) could be produced. As the density can be quite high, perhaps 10²¹ cm⁻³, because of the enormous magnetic pressure (1 Mbar), the required confinement time is relatively small, 10⁻⁷ s. Cru-

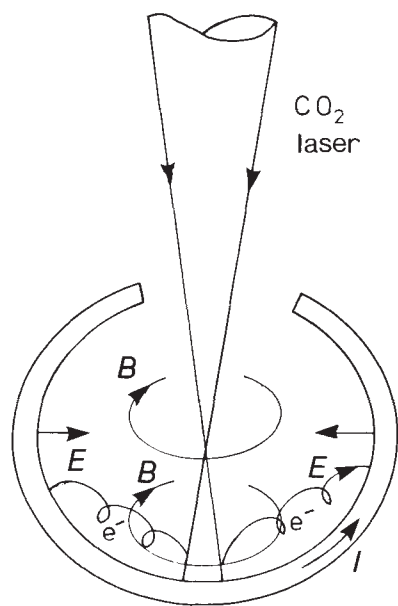
cial to the success of the scheme is the requirement that the magnetic field persists for some time after the laser pulse, and that the central hot plasma is able to refuel itself by ablation of cold material from the inside walls.

Unlike laser fusion there is no need to compress the fuel with the attendant problems of symmetry and fluid instability. There might be magneto hydrodynamic instabilities in the magnetic field that is used to sustain a large temperature difference between the reacting fuel and the walls, but the authors suggest that the system is stable against these effects because of its uniform total pressure.

In support of their rather radical proposals the authors present results obtained with a modest CO₂ laser at Osaka University. Using a 1.5-ns laser pulse focused inside hollow plastic spheres up to 3 mm in diameter they show that the hot plasma lifetime exceeded the duration of the laser pulse by up to 10 times.

It is too early to assess the prospects of this approach to fusion, as many details of the scheme have not been analysed and experiments are at an early stage. But the work is indicative of substantial Japanese involvement in all branches of fusion research and shows their increasing commitment to this research at a time when the United States and Europe are reducing their fusion programmes. □

R.G. Evans and M.H. Key are at the Rutherford Appleton Laboratory, Didcot OX11 0QX, UK.



Schematic of Japanese hybrid laser-magnetic confinement target showing *B*, *E*, electron trajectory *e*⁻ and current flow *I*.

machine is limited not only by the electrical power needed to drive the magnets but also by the sheer material strength needed to contain the magnetic pressure, which is proportional to *B*². As the contained plasma pressure cannot exceed the magnetic pressure, this limits the plasma density in a magnetic machine and requires that it contains the plasma for a time of the order of one second for an efficient fusion reaction.

Much effort, particularly at the Los

Immunology

Antigen binding and T cells

from N.A. Mitchison

LYMPHOCYTE recognition of antigens is a crucial feature of the immune response and occurs by means of receptor molecules. Although these receptors have fundamentally the same structure on the two major types of lymphocytes they work in very different ways. B cells, which use immunoglobulin molecules as their receptors, can bind antigen without the intervention of other molecules. This is not the case for T cells which normally bind antigen only when it is associated with a glycoprotein encoded by the major histocompatibility complex (MHC) on the surface of another cell. The evolutionary reasons for this 'dual recognition' are by now well understood. It provides a mechanism for ensuring that T cells come into action only when they touch another cell, which is important because they regulate or kill other cells through cell-to-cell contacts or at short range, unlike B cells whose secretions act far off. Two papers in this issue, Watts, T.H. *et al.* (*Nature* **320**, 179; 1986) and Ashwell, J.O. & Schwartz,

R.H. (*Nature*, **320**, 176; 1986) take us a step forward in understanding how T cells recognize antigen.

Any mechanism of this sort needs to be adjusted so that free antigen cannot block the T-cell receptor and MHC molecules do not activate on their own. Part of the adjustment seems to be pre-programmed in the receptor structure, but also selection of suitable receptors occurs in the thymus — hence 'restriction', the ability of T cells to recognize antigen only in the context of the MHC molecules to which they were exposed in the thymus.

Clearly an arrangement of two entirely separate receptors would have difficulty in meeting these requirements and would in any case be incompatible with current knowledge of receptor molecular genetics, so other possibilities have been canvassed. The simplest is a termolecular complex, in which receptor, MHC molecule and antigen all interact with one another. It would still seem difficult to adjust the affinities so as to prevent

bimolecular binding of receptor to antigen or MHC molecule (remember that we are dealing not with single molecules in solution, but with membrane-bound arrays in which high avidity could easily build up). So a good guess is that binding of two of the molecules may induce a conformational change that permits binding of the other to take place.

The answers may come from experiments in which these molecules are freed of their membrane-binding domains by genetic engineering and allowed to interact in solution, as several groups are currently trying to do. Watts *et al.* use fluorescence to study energy transfer from a peptide antigen to an MHC molecule in an artificial lipid membrane, and find that the two bind together only in the presence of specific T cells. Ashwell and Schwartz use less direct methods to address a similar problem. In previous experiments, by reading out activation of T-cell clones, they have been able to measure independently the potency of antigen-MHC combinations (from the concentration of antigen needed) and avidity for a receptor (from the ability of cells bearing the receptor to compete with one another). In the present study they show that switching an MHC molecule can change potency without affecting the avidity, and attribute this to a change in the strength of binding between MHC molecule and antigen. Certainly no other interpretation seems at all obvious.

Does the fluorescence signalling result contradict either this interpretation of potency/avidity or the earlier claim of Babbitt *et al.* (*Nature* 317, 359; 1985) to have demonstrated direct binding of anti-

gen by MHC molecules? As Schwartz himself pointed out in these columns (*Nature News and Views* 317, 284; 1985), the observed affinity did seem rather high, and it would be hard to reconcile such a value with the failure of MHC molecules to bind antigen in the absence of T cells as observed here. The direct binding experiment needs to be repeated in a criss-cross, using two MHC molecules predicted to bind two peptides reciprocally. On the other hand the new pieces of evidence do not contradict one another, because the MHC-antigen binding postulated to account for the potency change could well be small enough to have been missed by fluorescence signalling; and as the signals showed that MHC molecules and antigen ended up in tight proximity, they support the possibility of some degree of specific binding.

The outcome is to have strengthened the evidence of direct binding between antigen and MHC molecule. This in turn will strengthen the belief that immune response (and immune suppression) effects of MHC genes can result from such binding, and hence that this is what controls disease susceptibility in human and animal populations. But it should be emphasized that a very strong case can also be made for other forms of immune response control by MHC molecules, for example, through self-tolerance perturbing the T-cell repertoire (Müllbacher, A. *et al.* *J. exp. Med.* 157, 1324; 1983), and that the forms of control are likely therefore to be multiple. □

N.A. Mitchison is Professor of Zoology at University College, Gower St, London WC1E 6BT, UK.

Palaeoceanography

Biomarkers for ancient climates

from Erwin Suess

ENVIRONMENTAL conditions of growth, particularly ocean temperature, are faithfully recorded by the coccolithophorids — widely distributed marine phytoplankton. The record lies in the relative abundance of long-chain alkenones — complex organic molecules of the lipid bilayer which control the fluidity of cell membranes and thus respond to temperature stress. Elsewhere in this issue (*Nature* 320, 129; 1986), S.C. Brassell and colleagues report that the chemical unsaturation of alkenones (the ratio of more saturated to less saturated ketone molecules among the isomers of C_{37} alkenones, here called the U_{37}^K index) in contemporary marine sediments varies directly with the present-day temperature of the ocean surface over the sites from which sediments and some suspended-matter samples were collected. These samples span wide latitudes

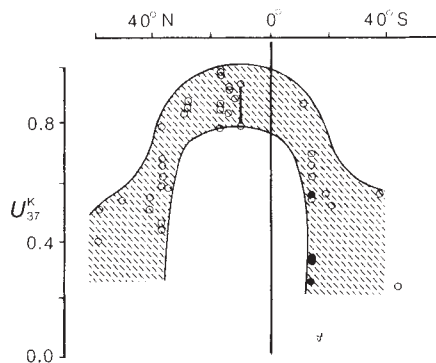
on both sides of the equator and bracket most of the global spectrum of sea-surface temperatures (see figure).

The authors show that the abundance of these molecular marker compounds in ancient sediments of the tropical Atlantic Ocean also varies with the expected ancient ocean climate. Glacial to interglacial fluctuations in climate, recorded by global oxygen isotope variations in calcareous shells of near-surface dwelling foraminifers, correspond to synchronous changes in C_{37} alkenone unsaturation indices. Such clear correlation provides a powerful new stratigraphic tool even in conditions where the temperature calibration is more tricky.

In the new approach, there are certain similarities to conventional oxygen-isotope stratigraphy. Using the latter, $\delta^{18}O$ of foraminifers was thought at first

to be sensitive mainly to temperature, but it later turned out to be a stratigraphic parameter recording a mixed signal of temperature and ocean volume change. The new report raises the intriguing prospect that a combination of the C_{37} alkenone unsaturation index with $\delta^{18}O$ data will eventually allow the portion of the oxygen-isotope signal caused by temperature to be separated from that related to ocean volume change.

As with most novel approaches, molecular stratigraphy also faces old problems. The deteriorating correlation between U_{37}^K indices and oxygen-isotope signals before 270,000 years ago, plus the lack of agreement in sediments more than 550,000 years old, provide clear evidence that conditions in the contemporary and ancient oceans are not analogous. One may surmise that a change in source organisms occurred at these times and that alkenone undersaturation may not be synthesized to the same degree by different species under the same environmental conditions. This hinders accurate temperature calibration more than stratigraphic correlation.



Alkenone unsaturation index for contemporary marine sediments and some sediment trap material from different latitudes. The distribution envelope corresponds to the pattern of global sea-surface temperature. Unsaturation index, $U_{37}^K = [C_{37:2}]/[C_{37:2} + C_{37:3}]$, where C_{37} denotes the total number of carbon atoms and :2 and :3 the number of double bonds per molecule. The greater the proportion of $C_{37:2}$ molecules, the higher the unsaturation index that corresponds to higher water temperatures. ○, Sediments; ● sediment traps.

The search for palaeoceanographic biomarkers — in the sense that earth scientists use index fossils, assemblages and ecological data of organic skeletons — appears to equal in difficulty that of finding needles in a haystack. There are an almost infinite number of known and unknown natural organic compounds that are potential biomarkers and which have to be explored and catalogued. Often they occur in minute quantities in sediments overwhelmed by other more abundant compounds that are of less diagnostic value. Recently, improved analytical techniques using computerized capillary gas chromatography and mass spectrometry have greatly improved the

chances of finding organic compounds diagnostic for either particular organisms or specific environmental conditions of growth (Mackenzie, A.S. *et al. Science* **217**, 491; 1982).

Evaluating the sources of biomarkers, their physiological functions in key organisms and their chemical stability over geological time will be major tests for the success of organic molecular stratigraphy. Detailed surveys of organisms are only one step towards the identification of sources. For example, dinoflagellate algae contain a unique sterol molecule, dinosterol, which also survives burial (Boon, J.J. *et al. Nature* **277**, 125; 1979). These marine algae attain peculiarly high blooms under particular oceanographic conditions during coastal upwelling. Dinosterol marks the presence of algal residues in marine sediments that might otherwise be devoid of fossils. Similarly, proteins of calcareous and siliceous marine plankton contain diagnostic assemblages of amino acids (Mueller, P.J., Suess, E. & Ungerer, C.A. *Deep-Sea Res.* in the press). Acidic amino acids dominate the proteins that secrete calcareous skeletons; basic amino acids dominate those secreting siliceous skeletons. Again, such organic compounds are the imprint of the biosphere in contemporary sediments.

A third type of organic molecular marker consists of those compounds diagnostic for certain environmental conditions such as temperature or pressure (deLong, E.F. & Yayanos, A.A. *Science* **228**, 1101; 1985). Organisms which adapt their cellular composition of phospholipids to temperature changes have long been known to biologists. Such organisms increase the proportion of unsaturated fatty acids, which have lower melting points and greater molecular flexibility than saturated isomers at lower temperatures. There is much effort now being made to identify the saturation indices of fatty acids of cell membrane phospholipids in marine microorganisms and benthic macroorganisms. The giant benthic communities at hydrothermal vents, which thrive at spectacularly high temperatures, are favourite targets for this research.

There is a problem in the poor chemical stability of amino acids and these lipid compounds. Their usefulness as biomarkers is restricted largely to young sediments. More complex and larger molecular structures appear more resistant to biodegradation and physico-chemical diagenesis, and thus have greater potential for survival in the geological record. The success of Brassel *et al.* in studying biomarkers lies in resolving the complexity and diversity of the large organic molecular structures.

A collaborative effort is now needed to understand the mechanism of stabilization of biomarkers in the geological record to exploit them fully, in particular if dissolu-

tion has severely decimated the conventional skeletal carriers of palaeoceanographic information in sediments, as in the vast red clay deposits of the central Pacific Ocean, which contain few or no microfossils. Fortunately, complex organic molecules may be stabilized and survive diagenesis by sorption to clays. It is of high priority to intensify the search for biomarkers in those sediments which are devoid of microfossils. Encouraged by the

new results of Brassel *et al.*, such efforts could unlock imprints of the past biosphere and environmental conditions of growth that are normally hidden from palaeoceanographic interpretation. □

Erwin Suess is visiting the Geological Institute, University of Kiel, 2300 Kiel, FRG, and is permanently at the College of Oceanography, Oregon State University, Corvallis, Oregon 97331, USA.

Behaviour rhythms

A *Drosophila* 'clock' protein?

from John Merriam

THOSE of us who have experienced jet lag are aware that biological rhythms, including daily, or 'circadian', rhythms affect our behaviour. There are thus commercial as well as academic benefits to be gained from understanding the biology of clocks and the gene products that are essential for their activity. One such gene, the *period* (*per*) locus of the fruitfly *Drosophila*, has been extensively studied, most recently by molecular cloning¹. Now Jackson *et al.* on page 185 of this issue² describe this potential clock gene, whose sequence predicts the amino-acid sequence of an unexpected protein which seems to be a proteoglycan. This is the first biochemical step in the identification of the biological clock, and helps to clarify a previously confused field.

Discovery of the *per* locus was one of the more remarkable sagas in applying genetics to biology³. The first three 'rhythm' mutants discovered were all alleles of the *per* locus, although they exhibited different phenotypes. The *per*-bearing flies have a long period rhythm, *per*^L flies a short period and *per*^S flies no circadian rhythm. Thus, in one step both the existence and the non-essential nature of an endogenous clock, as opposed to externally generated behaviour rhythms, was demonstrated. The *per* length mutants also affect the ability of flies to reset their clocks. These aspects of the *per* locus suggest that its gene product is close to, perhaps part of, the clock mechanism.

The location of the *per* gene mutants on the *Drosophila* X chromosome was established by standard recombination and cytogenetic techniques to be at the salivary chromosome bands 3B1-2 (see figure). By coincidence this also happens to be one of the most heavily studied genetic regions in the species: several chromosome deletions (termed deficiencies, or *Df*) which remove this region also remove the normal *per* function. Among the most useful *per*^S-deleted chromosomes are two that overlap only around the *per* gene. Female heterozygotes with *Df*(1)62d18 on one X chromosome and *Df*(1)64j4 on the other

are viable, although *per*^S in phenotype. The overlapping endpoints of the two deletions are markers of the left and right sides of the chromosome region responsible for the *per* function activity.

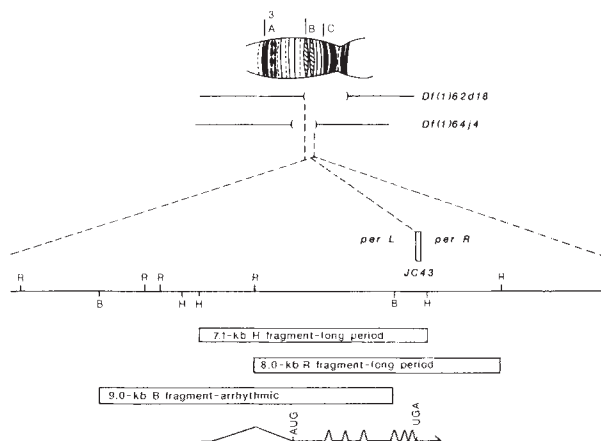
Another chromosome rearrangement with a breakpoint at 3B1-2, the translocation *JC43*, causes flies to exhibit an abnormal circadian rhythm which has a long period. This is consistent with the break causing a weak or 'hypomorphic' mutant allele, and thus represents another molecular marker of the *per* gene. Because the *JC43* translocation separates the X chromosomes into a left and a right part, the residual *per* activity can be mapped and is shown to be carried by the left fragment.

Working independently, groups at Rockefeller⁴ and at Brandeis⁵ universities cloned chromosome bands 3B1-2 and identified these left and right deficiency endpoints. They turn out to be separated by about 15 kilobases (kb); therefore the *per* gene is thought to be located within this region of sequence — but where? The *JC43* break divides the 15 kb into a left (*per*-L) and a right region (*per*-R). Several transcript-coding regions are found within these regions. Most attention has focused on a 4.5-kb head-specific transcript originating largely from *per*-L, but Reddy *et al.*⁵ have also identified a 1-kb transcript from the same region of *per*-L and a 0.9-kb transcript originating just to the right of the *Hind*III restriction site in *per*-R. (Restriction enzymes such as *Hind*III are used to cut DNA into fragments at specific regions of the sequence, as shown in the figure.) This last transcript is novel in that its abundance shows rhythmic behaviour and is reduced in the *per*^S mutants.

The simplest assumption is that the 4.5-

1. Rosbash, M. & Hall, J.C. *Cell* **43**, 3 (1985).
2. Jackson, F.R. *et al. Nature* **320**, 185 (1986).
3. Konopka, R.J. & Benzer, S. *Proc. natn. Acad. Sci. U.S.A.* **68**, 2112 (1971).
4. Bargiello, T.A. & Young, M.W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2142 (1974).
5. Reddy, P. *et al. Cell* **38**, 701 (1984).
6. Bargiello, T.A., Jackson, F.R. & Young, M.W. *Nature* **312**, 752 (1984).
7. Zehring, W.A. *et al. Cell* **39**, 369 (1984).

The salivary chromosome bands 3B1-2 (top of figure) contains break-points for deleted (*Df*) and translocated chromosomes (*JC43*). Location of these breakpoints on the restriction map (middle) of about 15 kb of cloned DNA from this region permits the identification of restriction fragments and a transcript (bottom) thought to represent the *per* gene product. R, *EcoRI*; B, *BamH*; H, *HindIII*.



kb transcript (or the 1-kb transcript) alone represent the *per* gene product. What that does not explain is why *per-L* (*JC43* left fragment) does not restore completely normal circadian behaviour, or indeed why the adjacent 0.9-kb transcript normally shows rhythmic behaviour. One scenario advanced by the Brandeis group is that the 0.9-kb transcript is also part of the clock. In this view both the 4.6- and 0.9-kb transcripts are necessary for normal clock activity.

DNA transformation experiments with cloned portions of the 15-kb interval have not completely settled this issue. In one case a 7.1-kb *HindIII* fragment, largely originating from *per-L* (see figure), when transformed, restored weak long-period activity⁶. In another case an 8.0-kb *EcoRI* fragment also restored weak long-period activity but a 9.0-kb *BamH* fragment from *per-L* did not restore any rhythmic behaviour⁷. The inability of any of these fragments completely to restore wild-type activity could mean either that the *per* gene is complex and includes several components or that none of these fragments, including the *per-L* region of *JC43*, completely encodes a rather long transcript.

On the basis of transcript mapping and DNA sequencing of the more than 7,000 bases in this interval it now appears that the latter alternative is more likely. Starting with a complementary DNA clone hybridizing to the 1.5-kb *BamH/HindIII* interval of *per-L*, Jackson *et al.*² succeeded in placing the transcript map onto the DNA restriction map. The transcript they describe includes eight exons originating over a slightly longer region than the 7.1-kb *HindIII* fragment previously used in transformation. The sequence is consistent with a single open reading frame with an AUG start codon in exon 2 and a UGA stop codon in exon 8. Of significance is the finding that none of the fragments used in transformation, or the *per-L* region of *JC43*, could be expected to encode a completely normal full-length transcript. For example, the *JC43* breakpoint, thought to fall about 170 base pairs upstream of the poly(A) addition site, is within the 502 base pairs forming the downstream un-

translated part of the messenger RNA. This strengthens the view that the 4.5-kb transcript alone forms the *per* gene. With this knowledge we can expect that transformation with a suitably longer fragment will completely restore wild-type rhythms. Understanding of the DNA alterations associated with the *per*¹, *per*² and *per*⁰ mutations, which should soon be forthcoming, will further pinpoint the gene. But none of this explains the behaviour or the role of the 0.9-kb transcript.

The predicted translation product of the open reading frame determined from the

per-L DNA sequence is a polypeptide of 1,127 amino acids. The amino-acid composition is somewhat unusual in that nearly half (47 per cent) comprises the amino acids serine, threonine, glycine, alanine and proline. Polyglycine and polyalanine repeats of up to 17 residues long are found at several positions. The putative *per* protein includes a potential site for phosphorylation by a cyclic AMP-dependent protein kinase. The amino-acid sequence at this site is homologous with the phosphorylation site in the calcium regulatory protein troponin I isolated from rat skeletal muscle. Other homologies with sequences found in the Dayhoff and Doolittle databases include the core protein of a rat chondroitin sulphate proteoglycan, predominantly in the alternating Ser-Gly residue series. These series, and the Thr-Gly repeats, may serve as sites for glycosylation. Nobody has speculated on the role of glycosylation sites in clock function, but site-directed mutagenesis techniques that alter the numbers or locations of these sites should permit the experimental test of their significance. □

John Merriam is Associate Professor of Genetics in the Biology Department, University of California, Los Angeles, California 90024, USA.

Visual cortex

Cholinergic input and plasticity?

from A.M. Sillito

THERE is a growing controversy in developmental neurobiology centred on the role of noradrenaline in the plasticity of the visual cortex of young mammals. One group maintains that depletion of noradrenaline results in a loss of plasticity, which can be restored by microperfusion (see refs 1,2), while others claim that there is no correlation between cortical noradrenaline levels and the presence or absence of plasticity in early life³⁻⁵. Now, on page 172 of this issue⁶, Mark Bear and Wolf Singer offer at least a partial resolution to the controversy.

In the so-called 'critical period' during the first three months of life, the visual cortex of cat or monkey shows a remarkable capability to reorganize its patterns of functional connectivity. The experimental model generally used to demonstrate this is that of monocular deprivation. Although both eyes normally influence an equal proportion of cells in the visual cortex, monocular deprivation (produced by suturing closed the lids of one eye) in the critical period results in an almost complete takeover of the visual cortex by the non-deprived eye⁷⁻⁹, concomitant with a corresponding loss in the capability of the deprived eye to drive cortical cells. These processes involve a major change in the

cortical distribution of the thalamic fibres that relay the inputs from the eyes and occurs only within the critical period. Even a few days of monocular deprivation can precipitate a major change in cortical organization⁹, whereas in the mature animal extended periods of deprivation over many months have no effect. The relatively sharp transition in the capability of the cortex to reorganize its patterns of connectivity has provoked considerable interest.

The involvement of noradrenaline was suggested by the imaginative experiments of Kasamatsu and Pettigrew^{1,2}, who used 6-hydroxydopamine (6-OHDA), a neurotoxin that is apparently selective for catecholaminergic terminals, to deplete cortical noradrenaline levels either by intraventricular injection or direct microperfusion of the cortex. Both procedures were found to block plasticity within the critical period — that is, monocular deprivation apparently fails to cause any redistribution of the pattern of influence of the two eyes over the population of visual cortical cells studied. Subsequent microperfusion with noradrenaline was reported to render the cortex again susceptible to monocular deprivation. These data broadly support the view that

the noradrenergic input to the cortex exerts an important control over plasticity.

There is general agreement that micro-perfusion of the visual cortex with 6-OHDA blocks plasticity in the critical period^{10,11}. But if cortical noradrenaline levels are depleted to comparable levels by any one of a range of other procedures, plasticity is unaffected. These include neonatal injection of 6-OHDA; electrolyte or 6-OHDA lesions to the dorsal noradrenergic bundle carrying fibres from the locus coeruleus; and intraventricular injection of another neurotoxin, DSP-4, also with fairly selective effects on noradrenergic fibres^{3,5,12}. Furthermore it now seems difficult to produce a convincing loss of plasticity by intraventricular injection of 6-OHDA. Taken together, these data make it hard to maintain the view that the noradrenergic input to the cortex is responsible for plasticity, but the data of Kasamatsu and colleagues are difficult to explain if the case for noradrenaline is dismissed^{2,13}.

The new report by Bear and Singer⁶ confirms that lesions affecting the noradrenergic input to the visual cortex have no effect on plasticity, but that there is a loss of plasticity when the noradrenergic lesions are paired with lesions to the cholinergic input. The authors also show that 6-OHDA apparently blocks the facilitatory effects of the neurotransmitter acetylcholine on the visual responses of cortical cells, which implies that 6-OHDA is effective because it causes both a depletion of noradrenaline and blocks the action of the cholinergic input. (Destruction of the cholinergic input alone does not block the plasticity.)

How does the new evidence relate to plasticity within the critical period? The visual cortex receives a series of non-specific modulatory inputs, most notably the noradrenergic input from the locus coeruleus, a cholinergic input from the magnocellular nuclei of the basal forebrain and a serotonergic input from the raphe nuclei. In the normal adult cortex these inputs influence potassium channels, cyclic AMP, cortical metabolism and possibly blood flow¹⁴⁻¹⁹. It is not easy to determine which of these effects in the adult cortex are relevant to plasticity in the critical period of young mammals. The most parsimonious explanation is that any procedure that severely disrupts the normal functioning of the cortex will block the processes underlying plasticity. In support of this view, it is found that micro-perfusion of the cortex with glutamate (which probably disrupts cortical activity either by direct excitatory action on cortical cells or as a consequence of neurotoxin effects) also blocks plasticity²⁰.

Thus, the debate still turns on one group of observations indicating that plasticity can be partially restored by micro-perfusion with noradrenaline or electrical stimulation of the locus coeruleus^{1,2}. If, as

Bear and Singer indicate, plasticity is disrupted only when both the cholinergic and the noradrenergic inputs to the cortex are blocked, it seems that either of these inputs can compensate for the loss of the other and that future efforts should be directed to some process that is influenced by both inputs.

There is evidence that both acetylcholine and noradrenaline can enhance the responsiveness of cortical cells to an excitatory input by inhibiting the potassium currents that underlie the after-hyperpolarization built up during a burst of action potentials^{16,18,19}. Acetylcholine seems to affect both a slow voltage-dependent potassium channel, as recently discussed in these columns¹⁸, and a slow calcium-activated potassium channel, whereas noradrenaline seems only to affect the calcium-activated channel. If the potassium-channel effects are important, one implication is that they raise the excitatory response to a given visual input above some critical threshold required for plasticity to occur. But acetylcholine exerts a much more potent and widespread facilitation of stimulus-driven responses in the visual cortex than noradrenaline, hence the cholinergic system should be more important to plasticity than the noradrenergic. The data of Bear and Singer do not suggest that damage to the cholinergic system has a significantly greater effect than the noradrenergic, so the action on potassium channels is unlikely to be the critical factor. Nonetheless it would be interesting to ascertain whether micro-perfusion experiments, with the aim to enhance or mimic the cholinergic influ-

ence, would restore plasticity in the way that is claimed for noradrenaline.

The potential role of the serotonergic input should not be neglected¹⁷ — any two of the modulatory influences may be able to sustain the conditions necessary for plasticity in the critical period. One might infer that the critical period is distinguished either by a heightened sensitivity of cortical processes (or some element of these) to the modulatory inputs or by a heightened output from these inputs. At present there is a notable lack of evidence for either possibility and therefore an urgent need for further corroboration of the conditions that can restore plasticity. □

1. Kasamatsu, T. & Pettigrew, J.D. *Science* **194**, 206 (1976).
2. Kasamatsu, T., Pettigrew, J.D. & Arey, M. *J. Neurophysiol.* **45**, 254 (1981).
3. Bear, M.F. & Daniels, J.D. *J. Neurosci.* **3**, 407 (1983).
4. Arien, J.P. *et al. C.R. Acad. Sc. Paris III* **295**, (1982).
5. Dae, N.W., Videen, T.O., Parkinson, D. & Rader, R.K. *J. Neurosci.* **9**, (1985).
6. Bear, M.F. & Singer, W. *Nature* **320**, 172 (1986).
7. Wiesel, T.N. & Hubel, D.H. *J. Neurophysiol.* **26**, 1303 (1963).
8. Hubel, D.H. & Wiesel, T.N. *J. Physiol., Lond.* **206**, 419 (1970).
9. Sherman, S.M. & Spear, P.D. *Physiol. Rev.* **62**, 738; (1982).
10. Paradiso, M.A., Bear, M.F. & Daniels, J.D. *Expl Brain Res.* **51**, 413 (1983).
11. Daw, N.W., Rader, R.K., Robertson, T.W. & Ariel, M.J. *Neurosci.* **3**, 907 (1983).
12. Trombley, P. *et al. J. Neurosci.* **6**, 266 (1986).
13. Kasamatsu, T., Watabe, K., Heggelund, P. & Scholler, E. *Neurosci. Res.* **2**, 365 (1985).
14. Magistretti, P.J. & Schoroderet, M. *J. Neurosci.* **5**, 362 (1985).
15. Videen, T.O., Daw, N.W. & Radar, R.K. *J. Neurosci.* **4**, 1607 (1984).
16. Madison, D.V. & Nicoll, R.A. *Nature* **299**, 636 (1982).
17. Foote, S.L. & Morrison, J.H. *J. Neurosci.* **4**, 2667 (1984).
18. Brown, D. *Nature News and Views* **319**, 358 (1986).
19. Sillito, A.M. & Kemp, J.A. *Brain Res.* **289**, 143 (1983).
20. Trombley, P. *et al. J. Neurosci.* **6**, 266 (1986).

A.M. Sillito is Professor of Physiology at University College, Cardiff CF1 1XL, UK.

African pre-history

Life in the Lake Victoria basin

from David W. Phillipson

THE East African interlacustrine region, in particular the western part, has occupied a key position in Africa's later prehistory. Here, at the headwaters of the White Nile are well-watered plains, 1,000–1,500 m above sea level, which support a dense human population practising, in the recent past, various traditional lifestyles. Today, the greater part of the area consists of savannah with patches of dense forest which, even within living memory, were more extensive than at present. Before human interference, this would have been the natural vegetation over the greater part of the region¹. Recent research, reported on page 164 of this issue², indicates that the process of forest clearance began earlier than previously thought, extending back over some five millennia. Further, it appears that preparation of land for cultivation was the prime cause of forest clearance, initially in the valleys,

which were clear by about 2,200 years ago (2,200 BP), and subsequently at progressively higher altitudes.

The interlacustrine region (see map) is a contact zone between diverse African agricultural traditions. Cereals are cultivated, notably the millets, including sorghum. But in much of the relatively low-lying country north and west of Lake Victoria, bananas are the staple food, accompanied by other vegetatively propagated food-plants, notably yams and sweet potatoes. These food crops arise from four distinct origins. The millets are indigenous African cereal crops which appear to have been brought under cultivation in what is now the Sahel or, in the case of finger millet (*Eleusine coracana*), in an area centred on Ethiopia and perhaps extending into northern Uganda³. Bananas, which grow at altitudes below 1,950 m, are generally believed to be of trans-Indian



The East African interlacustrine region bounded in the west by the Ruenzori mountains and equatorial forest. Scale bar, 300 km; shaded area, land over 1,500 m.

Ocean origin, introduced via the East African coast sometime in the first millennium AD. The yams and sweet potatoes are diverse, including not only African species of *Dioscorea* which may have been the first cultivated on the West African forest fringes, but also those of New World and trans-Indian Ocean origin.

Cattle occupy an important place in the traditional lifestyle of many of the region's inhabitants where this is not precluded by tsetse-fly infestation. Herding is a savannah activity that may be assumed to have begun following forest clearance.

Although the beginning of cultivation in the region may have been the reason for the onset of substantial clearance of the native forest, today the main reasons are the quest for firewood and the production of charcoal. Charcoal production may, however, have been a more significant early factor than is commonly realized, as recent archaeological investigations⁴ have revealed that this area, including the western Lake Victoria shore, saw the emergence of a sophisticated iron-smelting technology (2,500–2,000 BP) — the earliest in sub-equatorial Africa. Experimental reconstructions of the prehistoric smelting process⁵ have confirmed the substantial amounts of hardwood charcoal that were used; the long-term effect on forest cover is thus clear.

Archaeological investigations do not yet reveal a comprehensive picture of human settlement in this region throughout the Holocene. At Ishango on the western (Zaire) shore of Lake Edward, a specialized harpoon-fishing settlement⁶ has been dated to probably 11,000–6,000 BP. Elsewhere, presumed hunter-gatherers maintained a mainly microlithic technology.

Although the herding of domestic animals and possible plant cultivation had begun in the Lake Turkana basin of northern Kenya by approximately 4,500 BP⁷, and in the Rift Valley and adjacent highlands further to the south by about 3,000 BP⁸, there is no direct evidence that farming and pottery-making peoples were established west of Lake Victoria at this time. The earliest form of pottery recognized in the latter area is called Kanyore ware, named after an island site in the Kagera River, which was being made by about 2,500 BP⁹. It is associated with chipped stone stools rather than with iron, but the mode of subsistence of its makers remains unknown. The people who inhabited the sites where Kanyore ware has been found are a prime focus for future archaeological research in the Lake Victoria region.

By about 2,500–2,000 BP a distinct pottery tradition was widespread in the region, its northernmost occurrence being near Kabelega (Murchison) Falls. The tradition as a whole has been named after a site at Urewe in south-western Kenya, but several local sub-styles may be recognized¹⁰. There may have been some chronological overlap between the production of Kanyore and Urewe pottery, but at Gogo Falls east of Lake Victoria an Urewe deposit is stratified over one which contains Kanyore ware¹¹. It is at sites of this period in Rwanda that there is the first clear local evidence for food production¹² — teeth of goats (or sheep) and seeds of sorghum and finger millet. Urewe pottery is of particular interest in being associated with the smelting of metals, noted above. The Urewe settlement of the Lake Victor-

ia region is generally seen as a source from which derived the metal-using, mixed-farming lifestyle adopted by about 1,900–1,600 BP over much of the southern half of the continent¹³. This change to a food-producing economy is widely accepted as having accompanied the introduction of Bantu languages into southern Africa¹⁴.

It was probably about 1,200 BP that a distinct roulette-decorated pottery type first appears in the archaeological record of the region¹⁴; its connections are almost certainly with pastoral areas further to the north. It is more widespread than its predecessors, suggesting that more land was available for human settlement. The cereal cultivation and iron-smelting of Urewe peoples would, as the palynological evidence confirms, have led to significant reduction in the forest cover, permitting the eventual growth of herds of cattle. It remains for archaeology to demonstrate the human populations responsible for the earlier stages of forest clearance now suggested in the Lake Victoria region. □

1. Hamilton, A. *Vegetatio* **29**, 21 (1974).
2. Hamilton, A., Taylor, D. & Vogel, J.C. *Nature* **320**, 164 (1986).
3. Harlan, J.R., de Wet, J.M.J. & Stemler, A.B.L. (eds) *Origins of African Plant Domestication* (Mouton, The Hague, 1976).
4. Schmidt, P.R. *Tanzania Notes & Records* **84**–5, 77 (1980).
5. Schmidt, P.R. in *Proceedings of Panafrikan Congress of Prehistory & Quaternary Studies* (eds Leakey, R.E. & Ogot, B.A.) 335 (Leakey Memorial Institute, Nairobi, 1980).
6. de Hainzelin, J. *Les Fouilles d'Ishango* (Brussels, 1957).
7. Owen, R.B. *et al.*, *Nature* **298**, 523 (1982).
8. Ambrose, S.H. in *From Hunters to Farmers* (eds Clark, J.D. & Brandt, S.A.) 212 (University of California Press, Berkeley, 1984).
9. Soper, R.C. & Golden, B. *Azania* **4**, 15 (1969).
10. Van Noten, F. *Azania* **14**, 61 (1979).
11. Collet, D.P. & Robertshaw, P.T. *Azania* **15**, 133 (1980).
12. Van Noten, F. *Histoire Archéologique du Rwanda* (Musée Royal de l'Afrique Centrale, Tervuren, 1983).
13. Phillipson, D.W. *African Archaeology*. (Cambridge University Press, 1985).
14. Ehret, C. & Posnansky, M. (eds) *Archaeological and Linguistic Reconstruction of African History* (University of California Press, Berkeley, 1982).

David W. Phillipson is Curator of the University Museum of Archaeology and Anthropology, Downing St, Cambridge, CB2 3DZ.

100 years ago



The La France above Paris. Facsimile of an instantaneous photograph executed at the Observatory of Physical Astronomy, Meudon. From *Nature* **33** 422, 4 March 1886.

Ecology

Carnivore dominance and herbivore coexistence in Africa

from Jared M. Diamond

MOST concepts of temperate-zone ecology are derived from studies in Europe and North America, but are such concepts really appropriate to the temperate zone of southern Africa? This issue was debated at two meetings last year in South Africa: a symposium on competition and coexistence, held by the Zoological Society of South Africa at Pietermaritzburg (23–25 July, 1985); and a workshop on terrestrial ecology at Houw Hoek organized by the Foundation for Research Development (1–3 August, 1985). Two studies of large mammals discussed at these meetings illustrate some of the special problems of southern African ecology.

The first is an unintended natural experiment on a large scale (Hugh Berry, Etosha National Park: *Madoqua* 12, 242, 255, 1981; 13 151, 1982). The gravel pits excavated in Etosha during road construction produced entirely unexpected large-scale changes in the mammal community of the plains. The trophic relationships and size structure of this community are outlined in *a* in the figure. As a result of the seemingly modest disturbance there were crashes in the wildebeest and zebra populations; declines in cheetah, brown hyaena and eland; and increases in lions, spotted hyaenas, jackal and springbok.

Berry reconstructed the causal sequence of these changes as follows (*b* in the figure): rainwater standing in the alkaline gravel pits became a reservoir for anthrax bacteria, to which wildebeest and zebra are especially susceptible; and the resulting abundance of sick or dead prey more than offset the effects of the decline in healthy prey, causing a population surge of the two dominant and largest species of carnivores/scavengers, the lion and the spotted hyaena, which are immune to anthrax. The lions reduced the population of eland, which suffered little from anthrax but are a favourite alternative prey of lions. The lions and spotted hyaenas competitively suppressed the next largest species of cat and hyaenid, the cheetah and brown hyaena, respectively. Relief from competing cheetahs and brown hyaenas, together with the abundance

of carcasses, permitted the next-largest carnivore/scavenger, the black-backed jackal, to increase. Competitive release was also seen among herbivores — the wildebeest and zebra crashes permitted the smaller and more anthrax-resistant springbok (which also profited from the decline of cheetah, its major predator) to increase.

Berry's work illustrates the unique insights that careful study of the natural environment can yield. Ecologists will never be permitted to kill 25,000 wildebeest out of curiosity, but the same result produced by anthrax has illustrated how predation and competition control population sizes of mammals on the Etosha plains.

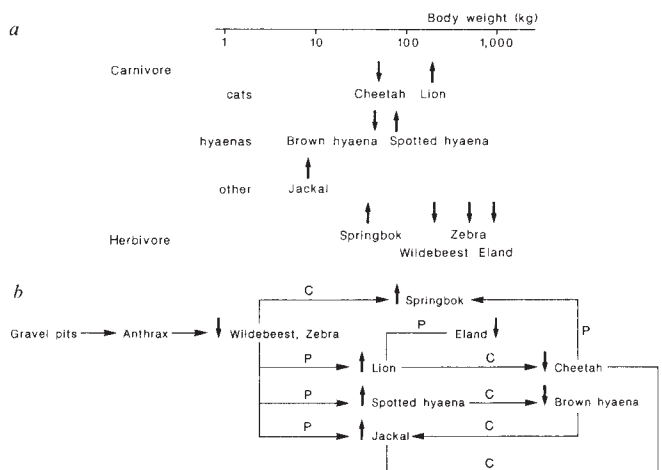
The second study concerned species coexistence among Africa's diverse large herbivorous mammals, which include 78 species of bovids. Several species often feed right next to each other and Richard Emslie (University of Witwatersrand) has examined the niche specializations that permit this coexistence among five large grazers in Umfolozi Game Reserve — white rhinoceros, buffalo, zebra, wildebeest and impala, adding a new level of detail to previous analyses.

After individual animals had finished grazing, Emslie searched their feeding sites for freshly defoliated surfaces, measuring the vegetation available within a 0.25 m² hoop and recording what had been eaten and how it had been eaten. Every plant in the hoop was examined and every tiller (the unit of a grass tuft) of the two key grass species *Themeda triandra* and *Panicum coloratum* was measured. He also monitored the availability of vegetation over an area of 33.25 km² by measuring within each 25-hectare block. This yielded data on resource selection over a hierarchy of levels, ranging from 25-hectare areas to parts of individual plants.

At the coarsest level, buffalo selected areas (25 hectares) with a high biomass of tall grass, whereas the other four species

chose areas of shorter grass. In the short grass areas, impala selected feeding patches dominated by *P. coloratum*, whereas wildebeest, often only a few metres away, sought patches with short *T. triandra*. Within feeding patches, the different species selected different tillers of the same grass species — the short erect tillers in the centre of *P. coloratum* plants grazed preferentially by wildebeest and white rhino, with their lawnmower-like wide muzzle and lips, and the longer, more prostrate tillers at the edge of tufts were eaten by the narrower-muzzled zebra and impala. Buffalo rejected tall *T. triandra* but selected medium-height tillers (40–75 cm), whereas wildebeest favoured short leafy tillers with no stem. Clearly, ecological studies that list only the species eaten omit considerations that are important to grazers.

Previous studies have revealed ecological differences among herbivores related to their differing body sizes and digestive anatomies (Bell, R. in *Animal Populations in Relation to Their Feeding Resources* ed. Watson, A. p. 111; British



a, Trophic relationships and size structure of mammals from the Etosha National Park. *b*, Inferred causal sequences of changes in the community. ↑, Increased abundance; ↓, decreased abundance; P, predator/prey interaction; C, competitive interaction.

Ecological Society, 1970 and Jarman, P. *Behaviour* 48, 215; 1974). Big animals cannot afford to be as selective about food as their related smaller species. On one hand, the non-ruminant zebra processes large quantities of low-quality plant matter quickly and inefficiently; on the other, ruminants process smaller quantities of higher-quality matter more slowly and efficiently. Emslie's studies now emphasize the importance of differing mouth anatomy as well. It is these differences that give rise to the familiar African scene of diverse large herbivores grazing next to each other, apparently in contempt of Gause's competitive exclusion principle but actually exploiting habitat heterogeneity at many scales. □

Jared M. Diamond is Professor of Physiology at the University of California Medical School, Los Angeles, California 90024, USA.

Corrigendum

The article by Jeremy Berg (*Nature* 319 264; 1986), which discussed the repeated domains in the TFIIA protein described by Miller *et al.* (*EMBO J.* 4 1609; 1986), should have mentioned that the repeating nature of the TFIIA sequence has been independently noted by R.S. Brown and coworkers (Brown, R.S., Sander, C. & Argos, P. *FEBS Lett.* 186, 271; 1985).

AIDS therapy by blocking CD4⁺ cells

SIR—In a recent *News and Views* article¹, Klatzmann and Montagnier discussed approaches to therapy of patients with acquired immune deficiency syndrome (AIDS), including immunosuppressive drug therapy, and, in particular, the case in which AIDS patients were treated with the drug cyclosporin A². Since AIDS results in an immune-deficient state, most efforts to restore immune function have involved immune-enhancing therapy³. However, the rationale for immunosuppressive drug therapy is that the development of AIDS involves a progression of immune disorders of which an early phase may be autoimmune⁴.

We suggest an alternative therapeutic approach based on the observation *in vitro* that the virus associated with AIDS, HTLV-III/LAV, is produced in activated but not resting T cells⁵. It seems logical to design therapy that would diminish the pool of CD4⁺ cells with the potential to produce virus efficiently. This could be accomplished by treating patients as early as possible in the course of the syndrome's development with either monoclonal anti-CD4 or anti-class II HLA reagents, either of which would block the activation of CD4⁺ T cells. The administration of CD4-specific antibodies would have the added advantage of interfering with the binding of HTLV-III/LAV to the CD4⁺ cell^{6,7}.

The mechanism by which HTLV-III/LAV causes the destruction of infected CD4⁺ T cells is unknown. If the virus is directly cytopathic, addition of CD4-specific reagents should prevent further infection of healthy cells, while permitting the infected cells to be lysed. It is likely that at least part of the destruction of infected CD4⁺ cells is due to cytotoxic T lymphocytes (CTL) that recognize HTLV-III/LAV antigenic determinants in association with self-HLA products and lyse infected cells⁴. In this case, treatment with antibodies that inhibit CD4⁺ helper T cell function should block further help provided by anti-HTLV-III/LAV-specific helper cells, but should permit the lysis of infected targets by CTL effectors that have already been activated.

From this perspective, it could be reasoned that immune-enhancing therapy such as administration of interleukin-2 (IL-2) would activate CD4⁺ cells and increase the number of T cells that would be susceptible to infection with the AIDS virus. Furthermore, since many HTLV-III/LAV infected individuals do not exhibit symptoms⁸, it appears that cofactors such as other infectious agents⁹, drugs¹⁰, or HLA allogenic stimulation^{11,12} contribute to the appearance of the syndrome. A common mechanism by which these various potential cofactors could operate is

by the activation of CD4⁺ T cells, thereby increasing the HTLV-III/LAV target cell pool size.

A therapeutic protocol that blocks CD4⁺ T cell activation would probably limit the number of CD4⁺ cells infected by HTLV-III/LAV and subsequently destroyed. However, because these antibodies function to prevent activation of helper T cells, they would be likely to induce immune deficiency, but one that is reversible upon cessation of treatment. Thus, therapy that blocks CD4⁺ T cell function differs from treatment with immunosuppressive drugs (which can have severe side effects and can be difficult to control) in that it is readily reversible. If therapy that blocks CD4⁺ T cells were to be attempted, it should be administered as early as possible in the development of AIDS, before depletion of CD4⁺ T cells has occurred, and probably with agents that have anti-retroviral activity^{13,14}.

ALFRED SINGER
GENE M. SHEARER

*Immunology Branch,
National Cancer Institute,
Bethesda, Maryland 20892, USA*

1. Klatzmann, D. & Montagnier, L. *Nature* **319**, 10 (1986).
2. Walgate, R. *Nature* **318**, 3 (1985).
3. Fauci, A.S. *et al. Ann. intern. Med.* **100**, 92 (1984).
4. Shearer, G.M. *Mt Sinai J. Med.* (in the press).
5. McDougal, J.S. *et al. J. Immun.* **135**, 3151 (1985).
6. Dalgleish, A.G. *et al. Nature* **312**, 763 (1984).
7. Klatzmann, D. *et al. Nature* **312**, 767 (1984).
8. Blattner, W.A. *et al. Ann. intern. Med.* **103**, 665 (1985).
9. Frederick, W.R. *et al. J. infect. Dis.* **152**, 162 (1985).
10. Marmot, M. *et al. Lancet* **i**, 1083 (1982).
11. Mavligit, G.M. *et al. J. Am. med. Ass.* **251**, 237 (1984).
12. Tung, K.S. *et al. J. Immun.* **135**, 3163 (1985).
13. Mitsuya, H. *et al. Science* **226**, 172 (1984).
14. Rozenbaum, W. *et al. Lancet* **i**, 450 (1985).

Serotherapy for AIDS and pre-AIDS syndrome

SIR—Acquired immune deficiency syndrome (AIDS) and the related pre-AIDS syndrome continue to spread¹ with as yet no effective treatment. Vaccines are a hope for the future, but it is generally acknowledged that vaccines are unlikely to prove useful in established AIDS. Recent suggestions of using methods of inhibition of gene expression to treat AIDS are theoretically attractive but also somewhat futuristic². There is another and more immediate approach to treating these diseases. McDougal *et al.*³ and others (including our unpublished observations) have shown that lymphadenopathy-associated virus/human T-lymphotropic virus type III (LAV/HTLV-III) infected T cells and T cell lines established from AIDS patients express the interleukin-2 (IL-2) receptor. This membrane receptor is an 'activation' marker of mitogen and antigen stimulated normal lymphocytes. The virus LAV/HTLV-III causing AIDS on gaining access to T lymphocytes via their CD4 receptors⁴ remains within a protected intracellular environment, avoiding the effects of neutralizing

antibodies. Such antibodies have been detected in the blood of AIDS and AIDS-risk patients^{5,6}. Present indications are that such neutralizing antibodies seem not to influence the course of the disease. This is probably due to the privileged intracellular location of the virus mainly within the helper/inducer T lymphocytes. This is indicated in the experiments of Robert-Guroff *et al.*⁶, showing that absorption with virus infected and uninfected H9 cells do not substantially affect neutralizing antibody titres, in contrast to cell-free virus preparations.

We suggest that the expression of the IL-2 receptor by the LAV/HTLV-III infected cells provides an extremely useful target (with a fairly high degree of tissue selectivity) for antibody-directed destruction of such cells. A high affinity anti-IL-2 receptor antibody (several monoclonals exist) selected from the appropriate species and of the appropriate isotype would be lytic-destructive for the LAV/HTLV-III infected target cells. The released virus would then be exposed to the neutralizing antibodies in blood. A good candidate for the anti-IL-2 receptor antibody would be a rat monoclonal of the IgG2b isotype. Such a monoclonal immunoglobulin with pre-defined specificity for membrane molecules of human and mouse T cells has already been shown *in vivo* and *in vitro* to be highly destructive to target lymphocytes. The antibody activates the recipient's own serum complement and operates in antibody-dependent cell cytotoxicity to give lysis of the targets⁷.

The AIDS and pre-AIDS patients' sera may contain enough neutralizing antibody to be effective against the exposed virus following the lysis of IL-2 receptor bearing infected targets. Alternatively, passive anti-LAV/HTLV-III human antibodies could also be infused to boost the patient's own antibody. Passive anti-LAV/HTLV-III antibody for administration can be obtained from pooled AIDS sera by ethyl alcohol fractionation which inactivates the virus⁸ without reducing antibody activity.

The combined serotherapeutic approach outlined above can now be put to the test using the recently developed *in vitro* model systems^{6,9}. Passive administration of antibodies is a time proven method of therapy, and the side effects are well documented and amenable to prophylaxis and treatment. Ideally this form of therapy should be introduced early after the diagnosis of AIDS and pre-AIDS syndrome, and possibly for patients shown to be LAV/HTLV-III antibody positive, with or without the secondary lymphocyte subset changes. Serotherapy could also be combined with drug treatment with agents such as suramin, which by itself in early clinical trials has proved to be ineffective in the long term, with marked toxic side effects. The early introduction of the ther-

apy described may ensure protection from infection of new T lymphocytes generated from the patient stem cells under thymic influence. Such cells could contribute to significant degrees of immune reconstitution of the individual.

H.F. SEWELL
F. WALKER

Department of Pathology,
University of Aberdeen,
Foresterhill, Aberdeen AB9 2XD, UK

1. Curran, J.W. *et al. Science* **229**, 1352 (1985).
2. Mariman, E.C.M. *Nature* **318**, 414 (1985).
3. McDougal, J.S. *et al. J. Immun.* **135**, 3151 (1985).
4. Dalgleish, A.G. *et al. Nature* **312**, 763 (1984).
5. Weiss, R.A. *et al. Nature* **316**, 69 (1985).
6. Robert-Guroff, M. *et al. Nature* **316**, 72 (1985).
7. Cobbold, S.P. *et al. Nature* **312**, 548 (1984).
8. Spire, B. *et al. Lancet* **ii**, 899 (1984).
9. Hoxie, J.A. *et al. Science* **220**, 1400 (1985).

Immunosuppressants in patients with AIDS

SIR—An interesting rationale for the treatment of acquired immune deficiency syndrome (AIDS) with the immunosuppressant cyclosporin A is based on the view that AIDS is an autoimmune disease¹. Immunosuppressive drugs would be designed to reduce virus replications and autoimmune processes triggered by retroviral envelope proteins that are bound to T4 molecules. Klatzmann and Montagnier state, however, that there is still no firm basis for such a therapeutic approach.

To this point, we are able to present a contribution that is experimentally corroborated by studies on neopterin excretion: measurement of neopterin levels provides a sensitive tool able to quantify activation of T-lymphocytes *in vivo*. Elevated neopterin levels are observed in virtually all diseases where activation of T-cells is apparent. *In vitro*, neopterin is released exclusively from macrophages in response to specific stimulation with γ -interferon (IFN- γ) secreted by activated T-cells².

Studies of neopterin excretion have demonstrated an activated cellular immune system in patients with AIDS, with AIDS related complex (ARC) and in risk-group members³⁻⁵. The frequency of elevated neopterin levels in risk groups is significantly correlated to the known AIDS risk factors: receptive anal intercourse, parenteral drug abuse and clotting factor substitution therapy in haemophiliacs. This elevation is observed also in recipients of multiple blood transfusions, infant children and in LAV/HTLV-III seronegative, apparently healthy risk-group members⁵. Although the activation of the cellular immune system seems to contrast with the clinical presentation of AIDS, the high serum levels of α_1 -thymosin and of acid-labile IFN- α in AIDS and ARC patients and elevated cytotoxic T-lymphocyte activity against allogeneic lymphoblasts in homosexual

men are consistent with activation⁶. These data argue for the view that T-cell-activation plays a crucial role in the development of AIDS. We conclude that conditions linked with T-cell activation are prerequisite for the development of AIDS. Permanent or multiple stimulations of the cellular immune system might represent the often discussed cofactor for progressive syndrome.

According to this concept, and as suggested by Klatzmann and Montagnier¹ and demonstrated *in vitro*⁷, the activation and proliferation of T4-cells will determine how much and how fast virus replication will ensue.

Consequently, T-cell activation should be strictly avoided particularly in LAV/HTLV-III seropositive individuals. Application of immunostimulative drugs such as interleukin-2 (IL-2), will exacerbate the course of disease by accelerating virus spreading and T4-depletion; the failure of IL-2 treatment is well known¹. Regimens preventing activation of T-cells and applications of immunosuppressive agents (such as cyclosporin A) that inhibit secretion of IL-2 from T-cells, should be tested with adequate caution.

Finally, the activation of T-cells by treatment can be readily assessed by measurement of serum or urinary neopterin concentrations which rise with immunostimulatory therapy and fall with immunosuppression.

Financial support by the Jubiläumsfonds der Österreichischen Nationalbank, P2591, is gratefully acknowledged.

A. HAUSEN
M.P. DIERICH
D. FUCHS
P. HENGSTER
G. REIBNEGGER
T. SCHULZ
E.R. WERNER
H. WACHTER

Institute of Medical Chemistry
and Biochemistry,

Institute of Hygiene,
University of Innsbruck,
A-6020 Innsbruck, Austria

1. Klatzmann, D. & Montagnier, L. *Nature* **319**, 10 (1986).
2. Huber, C. *et al. J. exp. Med.* **160**, 310 (1984).
3. Wachter, H. *et al. Hoppe Seyler's Z. physiol. Chem.* **364**, 1345 (1983).
4. Fuchs, D. *et al. Lancet* **ii**, 1130 (1985).
5. Wachter, H. *et al. Lancet* **i**, 97 (1986).
6. Tung, K.S.K. *et al. J. Immun.* **135**, 3163 (1985).
7. McDougal, J.S. *et al. J. Immun.* **135**, 3151 (1985).

Genetic prediction of cystic fibrosis

SIR—Brock and van Heyningen (*Nature* **319**, 184; 1986) have reason to be a little indignant about the way their current prenatal test for cystic fibrosis was ignored, but they are wrong to state that first trimester prenatal diagnosis by gene tracking with linked DNA probes suffers from a major problem due to the inability to confirm cystic fibrosis in the aborted

embryo. Their "cardinal requirement of prenatal testing, namely that some method of confirmation of diagnosis on the abortus be available" only really applies to tests based on empirical correlations between the level of some factor in amniotic fluid, for example, and fetal genotype (tests such as Brock's current prenatal cystic fibrosis test and his earlier important contribution, the α -fetoprotein test for open neural tube defects like spina bifida). This is not so for genetic prediction using linked probes.

The reliability of gene tracking using linked DNA probes depends on both exclusion of non-allelic genetic heterogeneity and the frequency of recombination between disease locus and marker(s). Careful linkage studies in a large number of families from different populations will allow exclusion of significant non-allelic heterogeneity and provide a mean recombination frequency. These studies are already well advanced. The recombination frequency combined with the prior genetic risk will allow a reliable prediction of the false positive and false negative rates for first trimester diagnosis; and the false negatives will also be observed after birth. Thus, the information required to assess the reliability of the prenatal test is available from studying living family members.

Our unit has probably had the greatest experience in Britain of using linked and gene specific probes for early prenatal diagnosis of haemophilia A, a disorder for which there is a reliable second trimester test by fetal blood sampling. Our experience shows that couples at risk understand the inherent errors in using linked probes, and yet almost always opt for the earlier test based on chorionic villus sampling; often followed, in the case of a negative result, by fetal blood sampling. We cannot over-emphasize the advantages of early diagnosis as perceived by the family. These include the psychological benefits of an early result and an earlier and safer therapeutic abortion, if they choose that option. A consequence seems to be a shorter gap before trying again if unlucky. Furthermore, there is the opportunity to keep the pregnancy private until the mother is decided.

MARCUS E. PEMBREY
SUE MALCOLM

Mothercare Unit of Paediatric Genetics,
Institute of Child Health,
University of London,
30 Guilford Street,
London WC1N 1EH, UK

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a rather technical character which are not provoked by articles or letters previously published.

Victorians abroad

David E. Allen

Explorers Extraordinary. By John Keay. *John Murray, London/Tarcher, Los Angeles: 1985. Pp.195. £10.95, \$7.95.*

The Travelling Naturalists. By Clare Lloyd. *Croom Helm/University of Washington Press: 1985. Pp.156. £13.95, \$19.95.*

Magnificent Voyagers. Edited by Herman J. Viola and Carolyn Margolis. *Smithsonian Institution Press: 1985. Pp.304. Hbk \$35; pbk \$17.50. To be distributed in Britain from May by Eurospan, hbk £37.95; pbk £18.95.*

EXPLORERS have always felt a need to find justifications for their efforts, and adding to scientific knowledge has long been one of the most popular of them. The Victorians, in particular, were fond of making out that the only reason why they hauled themselves up precipices or hacked their way through jungles was to make observations of unrecorded phenomena or to add to the tally of the Earth's known species. It was just not done to admit that other and less lofty motives lay behind their efforts: the physical challenge maybe, the thrill of the chase, or the glory of getting there first. A new literary genre came into being as a result: the scientific travel book. In addition to experiencing vicariously all the wonders and the perils of the untrodden wilds, the armchair reader wanted the reassurance that it had all been in a worthwhile cause, that the enjoyment and the suffering had had an underpinning of usefulness. And as the supply of scientists who had travelled interestingly was strictly limited, many who were barely scientific rushed in to fill the gap.

In his entertaining, light-hearted romp of a book — originally a series of radio talks — John Keay has disinterred a choice handful of the often rich and varied characters, a few of them frauds, most of them engagingly eccentric, from among whom the outer fringe of this literature acquired many of its inhabitants. Three with scientific pretensions feature among the seven whose exploits he recounts. The claim of one of them, H. Savage Landor, to have been contributing to science, however, is evidently to be taken no more seriously than the title *Everywhere* that he had the effrontery to give to the published narrative of his journeys. The other two, though, did make useful collections. Ludwig Leichhardt, one of a series of impossibly prickly Central Europeans who lend added colour to the history of the early exploration of Australia, had undoubted talents as a botanist and his eventual, still-unsolved disappearance was a definite loss to that discipline. For Mary Kingsley, niece of the author of the *Water Babies*, on the other hand, collecting was a sideline: it was to carry on her late father's researches into primitive religion that she braved the depths of the African rain-forest. Spe-

cializing in creatures that did not need to be taken with a gun, she brought back for the British Museum many valuable specimens of beetles and fish.

John Keay is a practised, professional writer, addressing with confidence what might almost be termed the professional



Perils of adventure: the Malolo incident of July 1840 in which Fijians attacked members of the US Exploring Expedition, killing Midshipman Wilkes Henry, nephew of the Expedition's leader.

reader. Clare Lloyd, by contrast, is a professional zoologist more obviously having to strain in writing for the less sophisticated audience on which she has set her sights. Her book, however, is reprehensibly mistitled: botanists and geologists, the other members of the trinity that ostensibly forms her subject, scarcely feature at all in its pages. It is hardly tactful of her, either, in a work whose readers will presumably be amateurs in the main, to let slip the suggestion here and there that genuine science can be done only by those who have received an academic training. Nor is the bibliography all that it should be, for much of the historical scene-setting is recognizably drawn from a source that she fails to mention.

Nevertheless *The Travelling Naturalists* is a book which, besides being read with pleasure, will prove of particular value for its helpful bringing together of accounts of the far-flung fieldwork of three late-Victorian naturalists: Howard Saunders

and Henry Seebohm, wealthy ornithologists who respectively followed their pursuit into the Andes and Siberia, and William Spotswood Green, who went a long way down as well as up, organizing the Royal Irish Academy's deep-sea dredging programme of 1885–1887 and notching up first ascents in Canada's Selkirk Mountains and in the Southern Alps. More familiar figures occupy more than half of the rest of the book: Mary Kingsley (again), Waterton, Bates, Speke and Franklin. While the tales of these certainly stand a re-telling, it is a pity the opportunity was missed to exhume rather more of the second division.

Bundling together biographies like this is one sure way of recapturing that age of the great pioneers: another is the portrait-in-depth of a single expedition. The sumptuous form in which the voyage of the *Beagle* was recently commemorated seems to have started a welcome fashion in this line of publishing; in the past year or two we have been treated to an even finer volume on the voyage of the *Endeavour*, complete with Sydney Parkinson's long-unpublished drawings, and now the Smithsonian Institution has surpassed even that with a yet more magnificent book on the Wilkes Expedition of 1838–1842. Just as the *Endeavour* volume was the product of a team of specialists drawn mainly from the ranks of the British Museum (Natural History), so their Smithsonian counterparts have turned out in force to provide us with a definitive account based on exhaustive historical and taxonomic research.

The United States Exploring Expedition (to give it its official title) was the largest voyage ever undertaken expressly for the purpose of scientific discovery. It was designed to demonstrate to the world that the youthful Republic was able to bring muscle to such work that was in no way inferior to that of the longer-established nations, and to the extent that the collections brought home proved as important as they were massive it can be said to have succeeded in its aim. The emphasis on national prestige, however, was very much a mixed blessing. On the one hand it wrung out of Congress — admittedly, only after many years of lobbying — funding on a scale generous enough to allow a staff of seven scientists to be recruited, foremost among them the 25-year-old James Dwight Dana, the effective founding father of American geology. It also decreed a certain showy expansiveness in the Expedition's itinerary, enabling a wide spread of territories to be visited and for collecting in these to extend in certain cases (most notably in Fiji) to spells of up to several months. The placing at its disposal of as many as half-a-dozen ships was a further piece of unheard-of lavishness. ►

On the other hand these were necessarily naval ships and as such quite unsuited to the scientists' requirements. The fact that the Navy held command, moreover, meant that surveying was accorded priority throughout. Worst of all, the Expedition's leader was high-handed and officious, even forbidding the conchologist member to take specimens below decks, thereby compelling him to restrict his collections severely.

Lieutenant Charles Wilkes, indeed, is both hero and villain of the story. But for his forcefulness and determination the Expedition would never have accomplished so much; yet, ever-conscious of the political purpose its research was meant to serve (not for nothing was he the great-nephew of the English politician of the same surname), he boxed in the taxonomists appointed to work up the collections with stultifying edicts. The reports, he insisted, must all be written by Americans in America — and, what is more, only in its capital city. Yet Washington at that date had no scientific library or reference collections of any consequence. Only after repeated protests did he consent to material being referred to specialists overseas.

Even when the scientific work was completed, its publication was ludicrously botched: the 19 volumes of reports and atlases in which it was embodied were limited to a print run of only 100 copies. The collections meanwhile were left without a home for all of 16 years, subjected to pilfering by Congressmen and other public figures who helped themselves to specimens as souvenirs. First offered to the Smithsonian in 1846, only to be rejected as too expensive to curate, they were finally consigned there in 1858 by a Congress bereft of further patience or ideas. Even then their fate was not necessarily happy. Most of the crustacean material perished while on loan to Chicago in the great fire of 1871. Symptomatically, too, the unpublished report on the fishes produced by Louis Agassiz was lost in the bowels of the museum for well over a century.

Despite all this, impressive additions to knowledge resulted. The zoological specimens alone turned out to contain nearly 2,000 new species. Dana's geological work, in particular, was to have an enduring influence, his observations in the Pacific anticipating the recent discoveries of plate tectonics. Many years were to pass, too, before the survey findings were superseded. Appropriately, indeed, even as late as the Second World War when the US marines came to invade Tarawa it was the chart made by the Wilkes Expedition that proved to be the only one available to guide them. □

David E. Allen, Lesney Cottage, Middle Road, Winchester, Hampshire SO22 5EJ, is a past President of the Society for the History of Natural History and author of The Naturalist in Britain: A Social History (Allen Lane, 1976).

Curiosities of crime in medicine

H.R.F. Keating

Surgeons at the Bailey: English Forensic Medicine to 1878. By Thomas Rogers Forbes. Yale University Press: 1986. Pp. 255. \$26, £20.

HISTORY has a terrible habit of lapsing into anecdote, and one never knows whether to cry or to laugh. This ambivalence applies just as firmly when the history is the history of science. Thomas Rogers Forbes, Professor Emeritus of Anatomy and advisor in medical memorabilia at Yale, has had the worthy academic notion of outlining a history of forensic medicine in Britain. He uses chiefly the Old Bailey Sessions Papers, records of trials at what is now the Central Criminal Court from 1684 up to an arbitrary finishing point of 1878.

These papers are one of the few sources we have for seeing what doctors or embryo medical scientists did and thought when they were brought into rough contact with the varied forms of human criminality. So far so good. Professor Forbes has delved into them with persistence and, with footnotes and bibliography (13 small-print pages), he has laid down what could have been a formidable account.

But, alas, there is a falling-away from the ideal. Fundamentally, the material for a proper account of the gradual introduction of scientific method into English criminology is not available in the Sessions Papers nor in such other accounts as Professor Forbes from time to time quotes. Furthermore he has chosen to present what evidence he has, not in one sweep, but piecemeal, with potted histories of advances in the detection of blunt instrument injuries, of asphyxia, of sexual offences, of poisoning, of insanity and of a score of other minor aspects.

So it is hard to get much sense of the development of forensic medicine as a whole. On the other side of the coin, the book is simultaneously a tremendous hotch-potch of curiosities and comicalities. Bedside reading for a month, if you can take the odd references to inflammation of the ileum and, indeed, to *peine forte et dure*, an early scientific approach to finding the truth by squashing a suspect till he talked.

The fact is that murder is as much risible as it is horrific, I suppose because for many of us the funny side is the only way we dare look on the possible prospect of our own violent demise. So we get the so-called Resurrection Men who supplied bodies for anatomical study by digging them up or on occasion doing in the living. Thus the poet Thomas Hood wrote,

Don't go and weep upon my grave
And think that there I be.
They haven't left an atom there
Of my anatomie.

We get, too, such curiosities as *cruenation*. *Cruenation* is the belief that if a murderer can be made to touch the corpse of his victim the wounds will at once bleed. "See", wrote Shakespeare, "dead Henry's wounds Open their congeal'd mouths and bleed afresh". And in 1658 a Major Strangeways was put to this test, passed it triumphantly — and then succumbed under the *peine forte*, thus sacrificing himself to prevent his property being forfeit to the Crown.

Other speculation-causing titbits include the fact that Addison's Disease is named after Dr Thomas Addison, who actually confused that malady with another (read the footnotes). Or some vicious textbook reviewing in the early 1800s in which accusations of bad grammar played as dominant a part as challenges to experimental conclusions.

But perhaps the most savourable pages come in the various accounts of attempts on the life of Queen Victoria. There was Edward Oxford, a mulatto pot-boy who, in 1840, shot twice at Her Majesty, probably with pistols without bullets, and, out of a contradictory plethora of medical evidence, was found insane. (Because he was tried for treason rather than attempted murder he was allowed Counsel without payment.) In the Bethlehem Hospital he learnt six languages, chess, carpentry, the violin and knitting, as well as meeting George Hadfield who, convinced he was King George III and wanting to commit suicide, had fired at the real George III and missed.

Robert Pate, who in 1850 merely whammed Her Majesty on the head with a walking-stick, inflicting only a cut, was luckier. Life being odder than fiction, he was arrested by the selfsame policeman who had seized M'Naughten, the error-prone assassin of Robert Peel's secretary and origin of the long-held M'Naughten Rules for defining insanity. Pate was found guilty and not insane — the Judge happened to disagree with the M'Naughten formulations — but was merely sentenced to seven years' transportation. A mad world, my masters — in every sense. □

H.R.F. Keating, 35 Northumberland Place, London W2 5AS, UK, is a crime novelist and creator of Inspector Ghote of the Bombay CID. His most recent book is Under a Monsoon Cloud (Hutchinson, 1986).

New in paperback

• *The Limits of Science*, by Peter Medawar, a characteristically short, incisive and elegant account of the place of science in the modern world. Publisher in Britain is Oxford University Press, price is £3.95. The hardback edition was reviewed by Lewis Thomas in *Nature* 312, 203 (1984).

Cells come of age

Dennis Bray

Annual Review of Cell Biology, Vol. 1 1985. Edited by George E. Palade, Bruce M. Alberts and James A. Spudis. *Annual Reviews Inc.*, 4139 El Camino Way, Palo Alto, California 94306, USA:1985. Pp.580. \$27 (North America), \$30 (elsewhere).

REVIEWING reviews can be a vertiginous business, especially in a field such as cell biology in which any single experimental observation may be caught up and reflected from a ramifying literature, like images in a Hall of Mirrors. On the red blood cell cytoskeleton, for example, you will have a choice of a dozen or more reviews written in the past four years. All have the same viewpoint (no spicy controversies here) and offer more or less the same litany of facts. It is reasonable to ask how useful they are: do we need more reviews?

Well, yes, we do, at least when they are as authoritative and critical as those published by Annual Reviews Inc. Founded in 1929, the company is a non-profit-making organization and produces volumes of reviews of scientific subjects at a remarkably low price. The directors, editors and authors all contribute their energies and talents for free, and in the interests of their science they are very conservative in starting new series. Beginning with the *Annual Review of Biochemistry* in 1931, the company now produces 27 volumes annually: a remarkably small number when one considers the range of subjects covered and the awesome increase in published literature in 55 years. New volumes are introduced only reluctantly, so that when one does appear it is a mark of recognition — a coming of age.

Some material in the new *Annual Review of Cell Biology* inevitably overlaps other disciplines. The articles on intermediate filaments, fibronectin, acetylcholine receptors and actin-binding proteins could as well have stayed in the parental *Annual Review of Biochemistry*. But the main part of the volume is distinctive in content and deals with a higher level of organization. Accounts of structures such as the brush border, the membrane cytoskeleton and microtubule organizing centres, and of organelles such as peroxisomes and Golgi apparatus, are complemented by refreshing, integrative reviews of the processes of protein targeting, endocytosis and membrane traffic, and the establishment of cell polarity in eukaryotes and prokaryotes.

The emphasis is still strongly molecular, and only one article — an excellent account of cell migration in the embryo by Thiery, Duband and Tucker — treats cells

as the sentient creatures we all know them to be. Obviously molecules are easier to study and generate more data. But cell biology and molecular biology are not the same (despite some confused statements to this effect in the preface), and it is a pity that such an influential volume does not have more on the lineage, differentiation, behaviour and tissue interactions of the eponymous cell.

Within the rigidly defined format of the *Annual Review* volume, many excellent features are to be found in the book. There are more illustrations than usual, adding greatly to the interest and comprehensibility of a subject that deals in organization and structure (line diagrams are probably more effective than halftones in such an economically produced volume). It is also pleasant to discover how readable

many of the articles are. Back in 1950, the first volume of the *Annual Review of Physical Chemistry* regaled its readers with sentences several pages long, containing lists of references and chemical names and one verb. We have come a long way since then. Many of the reviews in this new volume could be enjoyed by informed readers outside the immediate area of specialization, especially where authors have ventured to bring the facts together in a hypothesis. The marvellously complete account of mitosis and meiosis by Murray and Szostak creates calm in a tormented universe; right or wrong, intelligent models of this kind certainly make facts easier to remember. □

Dennis Bray is in the Medical Research Council's Cell Biophysics Unit, 26-29 Drury Lane, London WC2B 5RL, UK.

Scientific anatomy

Philip Gummert

About Science. By Barry Barnes. *Basil Blackwell*:1985. Pp.163. Hbk £19.50, \$24.95; pbk £6.50, \$8.95.

SCIENCE is the dominant form of cognitive knowledge in all modern societies. What counts as empirical knowledge in these societies is very close to being what scientists and associated experts allow so to count. How this state of affairs came about, and what it entails for both science and society, are the main themes of this book.

These themes are the bread and butter of the social studies of science, a field which developed largely from the 1960s, when there was a flurry of interest in the possibility of a "science of science". Although that goal has long been abandoned, we do now have a richer understanding of science as a social institution. Unfortunately, much of the relevant literature is written only for the *cognoscenti*. Barnes redresses the balance well with this extremely lucid book. His intended audience is undergraduates and sixth-formers who are studying natural science and wish to learn more about science as an activity. But the book is so well written that it should attract a much wider readership.

Barnes describes how science became central to society during the Industrial Revolution. This was not because there were large numbers of scientists, nor because of any misplaced belief in its usefulness. Rather, science was for the rising commercial and industrial classes what the Bible and the classics had been for the old landed gentry and aristocracy. As science developed so did specialization, and with it came social processes for authenticating scientific knowledge and granting admis-

sion to the body of acknowledged scientists. Within this new social institution, recognition by one's peers served the purpose that money served in the wider society: scientists sought recognition of their work, not as an end in itself, but because with it they could more readily advance their careers in the ways that they wanted.

The unequal distribution of recognition among the scientific community means that the word of some scientists counts more than that of others. This has the useful consequence that the authority of the most highly-regarded scientists acts as a filter for the mass of data which scientists encounter daily, and so allows them to survive without having constantly to go back to basics. This portrait of scientific activity contradicts, of course, the popular picture of scientists as always questioning everything.

The authority structures that arise within science extend into the wider society. This did not happen by accident. As part of their campaign in the nineteenth century for social acceptance, scientists fought the church for recognition as the authoritative voice on issues to do with natural phenomena. Their victory is now more or less complete, and this raises the question today of the possible political domination of society by experts. Barnes is sceptical of this argument. For him, society is dominated *through* science and technology rather than *by* them. He also disputes the proposition that, if only all citizens were properly informed about the issues of the day, rationally correct decisions would be made by simple aggregation of individual preferences. Such aggregation is neither simple nor even necessarily rational. The science of society remains as remote as the science of science. □

Philip Gummert is in the Department of Science and Technology Policy, University of Manchester, Manchester M13 9PL, UK.

At the threshold of consciousness

K.E. Webster

The Thalamus. By Edward G. Jones. Plenum: 1985. Pp.935. \$135, £114.09.

SOME 50 years ago, Sir Wilfrid Le Gros Clark described the thalamus as "at the threshold of consciousness". The phrase conveys vividly the peculiarity of an organization common to all vertebrates: the flow of sensory information to the cerebral cortex and basal ganglia (or their homologues) in the forebrain is, with few — and largely peculiar — exceptions, strikingly interrupted in the nuclei of the thalamus. And yet the effects of the destruction of the thalamus, even in human beings, are seemingly either obviously predictable (for example blindness) or spectacularly unpredictable (for example the changes in pain sensibility associated with the thalamic syndrome). This crude summary indicates, I hope, why the thalamus has fascinated so many distinguished neuroscientists for so long.

Nonetheless, this beautifully produced book is the first large-scale monograph on the thalamus to be published since the 1930s. Its scope is wide: it includes an introductory historical chapter — which lends an admirable perspective to the whole — and ranges from cytoarchitecture through synaptology, connectivity, physiology and histo-pharmacology, to ontogeny and comparative studies of the vertebrate thalamus. In general, the functional groups of nuclei are dealt with together. In many cases, when the sources of afferents and the targets of efferents are referred to, the discussion of the organization of the sources and targets is commendably extensive and well-researched: comparable in fact, to that on the subject matter *sensu strictu*. The reader can thus turn to the book as a useful source of references and informed discussion on, for example, the organization of the cerebral cortex.

The text is very readable, yet exhaustively referenced, and the numerous illustrations are of high quality. The book imparts a clear sense that the overall objective is to explain the significance of the vertebrate thalamus. That, in the last analysis, the account fails in this endeavour is a reflection not so much upon the author as upon the ultimately intractable nature of his subject matter. Professor Jones makes us well aware that, vast as our knowledge may now be, our understanding remains small: a truly all-embracing synthesis is as little possible now as it was 50 years ago.

Inevitably, in a book of this size and compass, there are details of interpreta-

tion with which one might disagree, or information at the omission of which one might complain. In my view such omissions are virtually always trivial, and represent but minor blemishes on a successful and accomplished work, while the contentious emphases and interpretations are, for me at least, the spice in a generous meal. This book is at once stimulating and useful: the earlier monographs of Le Gros Clark and of Walker have a worthy successor. □

K.E. Webster is Professor in the Department of Anatomy and Human Biology, King's College London (KQC), Strand, London WC2R 2LS, UK.

Brief encounter

Irwin Fridovich

Biochemistry of Dioxygen. By Lloyd L. Ingraham and Damon L. Meyer. Plenum: 1985. Pp.287. \$45, £39.13.

THIS volume, fourth in a series entitled *Biochemistry of the Elements*, might have had as its subtitle: *Once Over Lightly*. Dioxygen, its singlet states and reduction products, and the enzymes which produce and act upon them, has been under intense investigation for several decades and is the subject of a vast and rich literature. Unfortunately the authors have tried to cover it all in a small-format book of only 287 pages. They would have been better advised to have chosen some modest fraction of this broad field and to have tilled it more diligently. Probably the cause for the less than satisfactory outcome was faulty planning; if selenium deserved one volume (Vol. 2 in the series), then dioxygen should have commanded at least four volumes.

The collision between the enormity of the subject matter and the paucity of space has yielded a book which does not bear up well under close scrutiny. The results of

the past five years of research are almost entirely ignored; illuminating equations, chemical structures and diagrams are too few; typographical errors are numerous; explanations are sometimes so compressed that their import will be opaque to most readers; and the facts stated are sometimes erroneous. Some specific examples are required to give substance to these charges.

Thus, sulphite is prone to oxidation by a free radical chain reaction, but sulphite does not reduce O_2 to O_2^- . The iron-containing superoxide dismutase of *Escherichia coli* was once thought to reside in the periplasm, but this claim was withdrawn on the basis of subsequent work. Some evidence for the existence of an essential persulphide in xanthine oxidase was once presented, but more recent work identifies the essential sulphur as a terminal ligand of the molybdenum, which in turn is attached to a novel pterin prosthetic group called molybdopterin; this pterin, whose discovery is the most exciting development of the past decade of study of xanthine oxidase, is nowhere mentioned. D-Amino acid oxidase catalyses the oxidation of the nitroethane anion, not of ethyl nitrate. The chapter on catalases and peroxidases makes no reference to the bacterial catalases, such as the manganese-catalase, and the chapter on "Biological Iron Dioxygen Carriers" fails to mention the Bohr effect and the influence of 2,3-diphosphoglycerate on the oxygen-binding curve of haemoglobin. The cooperativity of dioxygen binding to haemoglobin is discussed but not its physiological significance.

There is much more, but this will suffice. One may sum up the book by saying that the authors tried too much and hence were doomed to fail. We must hope that another attempt by several authors, spread over several volumes, will one day do justice to this important subject. □

Irwin Fridovich is James B. Duke Professor of Biochemistry at Duke University Medical Center, Durham, North Carolina 27710, USA.

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Tin decline: an abandoned mine, near Lydford, Devon, UK. The picture is taken from the paperback edition of Landscape in Britain, with photographs by Charlie Waite and commentary by Adam Nicolson, to be published in Britain on 24 March by Pavilion.

Investigating the paranormal

from David F. Marks

Parascience has so far failed to produce a single repeatable finding and, until it does, will continue to be viewed as an incoherent collection of belief systems steeped in fantasy, illusion and error.

FEW fields of inquiry capture the attention of the public as much as the paranormal. Newspapers, books, films and television have all cashed in and promoted it. Yet after millennia of experience and more than a century of controlled investigation, since the founding in 1882 of the Society of Psychical Research, the paranormal remains as controversial as ever. While credence in extrasensory perception (ESP) and precognition is widespread, parapsychology has failed to produce a single repeatable demonstration. In the face of such a dearth of hard evidence, how can such widespread belief in the paranormal be accounted for?

The importance of rigorous analysis of the evidence for parascientific claims cannot be underestimated. The establishment of ESP could conceivably require a paradigm shift (in Kuhn's sense) of the most fundamental kind and our concepts of mind-brain relationships and consciousness would need radical alteration. Our whole approach to psychology as an empirical science, based as it is on the time-honoured assumption that perception can result only from sensory activity, would be brought into question.

The conventional response of many scientists to the paranormal is to ignore the evidence on *a priori* grounds, believing it to be of basically poor quality. This attitude is allied to the Humean stance that a lie is more probable than a miracle. Although this scepticism is certainly justified, it could be argued that such a blanket response is counterproductive. First, it is hardly scientific to reject a claim purely because of its *a priori* improbability. Second, a division is created between aligned groups of committed 'believers' and 'sceptics' and the resulting adversarial positions inhibit proper discourse and the possibility of an account which satisfies all parties. Third, it leaves the field open for undisciplined exploitation, which is irresponsible; there are many examples of financial loss, suffering and even death resulting from fraudulent paranormal claims (for example, the Jonestown massacre, psychic surgery, the Transcendental Meditation levitation programme, firewalking, scientology and other pseudo-scientific cults). For scientists passively to

ignore such developments is, to say the least, uncharitable.

The Committee for the Scientific Investigation of Claims of the Paranormal (CSICOP) was established in 1976 with the aim of increasing the quality of scientific investigations into the paranormal by constructive criticism and the exposure of invalid or fraudulent claims. Over this 10-year period, an inordinate amount of fraud, error and incompetence in paranormal investigations has been brought to light¹⁻¹¹. But pseudo-sciences are remarkably stable and tradition-bound; their presence on the edges of science can be expected indefinitely¹².

Areas of experimental psychology can shed light on the paranormal, especially the study of consciousness and cognition. Such investigation indicates that the many anomalies of putatively paranormal experience are an inevitable consequence of normal selective and constructive processes in perception, memory and imagery. I summarize here what seem to be the common assumptions on which claims of the paranormal are based.

Theoretical assumptions

Paranormal phenomena are negatively defined. A phenomenon is defined as paranormal only if it contravenes some fundamental and well-founded assumption of science. Hence, to establish an effect as paranormal, all possible 'normal' explanations must be shown to be invalid. Any paranormal claim thus also remains provisional; a normal explanation, not previously thought of, may be discovered at some future time.

Mysteriousness *per se* is a necessary but insufficient condition for adjudging an event as paranormal. There will always be limits to knowledge, so that new phenomena that initially appear anomalous will be given a natural explanation following systematically controlled observations. Bona fide paranormal effects, on the other hand, are supposed to contravene established assumptions as though from another order of existence and not simply for lack of explanation. 'Contranormal' would be a more precise technical term.

Examples of effects which until recently

were claimed to be paranormal but which can now be explained from within orthodox science include:

(1) Kirlian photography, the photographic recording of coronal discharges around living or non-living objects produced by high-voltage, (20–100 kV), high-frequency (75 kHz–3 MHz) electrical fields^{13,14}. Variations in the images of the corona can be explained in terms of normal physical factors such as moisture, pressure or distance, all of which influence circuit resistance.

(2) Fire-walking, if conducted briskly on hot materials of low thermal capacity and poor thermal conductivity, does not produce burns^{15,16}. The Leidenfrost effect created by an insulating layer of water or sweat may also reduce energy transfer to the surface of the body.

(3) Dowsing is based on sensory cues, expectancy effects and probability. Controlled trials fail to produce above-chance results^{8,17-19}.

(4) Psychic surgery, thought photography and metal bending all involve sleight-of-hand and can be duplicated by skilled magicians^{8,20-22}. The first differs from the others in respect of the associated false hopes and financial loss, but all three are fraudulent.

(5) 'Gellerized' watches^{21,22}, thought to be broken, are purportedly repaired by illusionist Uri Geller by 'psychic concentration'. In about 50% of cases, simply holding the watch in a clenched fist and shaking it provides a sufficient stimulus to free the mechanism²³.

(6) Astrology, graphology, tea-leaf and tarot card readings, the I Ching and other forms of divination are all types of 'cold reading' or 'sleight-of-mouth'²⁴. They depend upon ambiguous, wish-fulfilling and general advice, the use of prior or presented information and cues obtained by verbal 'fishing'. A strong feeling of personal validation often accompanies such readings. Various forms of 'mediumship' and 'psychometry' as practised by D. Collins and D. Stokes are also examples of cold reading.

In some cases, field observations can be checked under laboratory conditions and the sensori-motor features of the original performance reproduced using a delayed

control group of non-psychic subjects; for example, Geller's watch-starting procedure and ability to draw the contents of sealed and apparently opaque envelopes were matched by that of non-psychic controls²³ (Fig. 1). Clearly, the tendency to judge a mysterious event as paranormal in the absence of controlled observations can be quite misleading.

The most dramatic evidence for the paranormal has been based on either fraud or methodological error. Apparent frauds that have been uncovered include University of London mathematician S. G. Soal's manipulation of his recording sheets²⁵⁻²⁷, University of Utrecht Professor Tanhaeff's evidence on Croiset, the Dutch 'psychic' detective²⁸, and the description by C. Castaneda (University of California at Los Angeles) of the paranormal teachings of Don Juan²⁹. C. E. M. Hansel³ has provided a valuable review of the history of trickery, fraud and error in parapsychology. However, outright fraud is not the only vehicle in which the paranormal cause can travel, and it is a serious mistake to assume it is a necessary part of any paranormal investigation.

There are no theories to account for paranormal effects or their properties. There are some undesirable implications of this aspect. First, investigators are unable to conduct properly controlled experiments on the properties of psi phenomena because they have no idea what the relevant variables are. In particular, there is no procedure by which psi can be deliberately switched on or off, and so there is no possibility of examining the effects of psi on other variables. All that can be done is to establish whether a given performance in some particular set of circumstances differs from a baseline; if so, psi is assumed to be the cause.

Oddly, neither the subject nor the experimenter can state which of the successful trials in a psi experiment result purely from chance guessing and which are generated by psi, so that there is no basis for distinguishing between a high score in a psi experiment and a lucky run in a game of chance. Also, the persistence of psi investigators in the face of variable but mainly negative results could have a similar motivational basis to that of addicted gamblers; both show high resistance to extinction following variable ratio schedules of reinforcement³⁰.

A more fundamental problem with the paranormal's atheoretical status is that of untestability. Failure to observe a particular effect can be readily attributed to a host of *ad hoc*, hypothetical factors. Vivid imagination is no substitute for testability, however, and if *ad hoc* hypotheses are not independently testable, nor is the original claim³¹.

It has been stated that the participants in any experiment may unconsciously determine the results according to whether

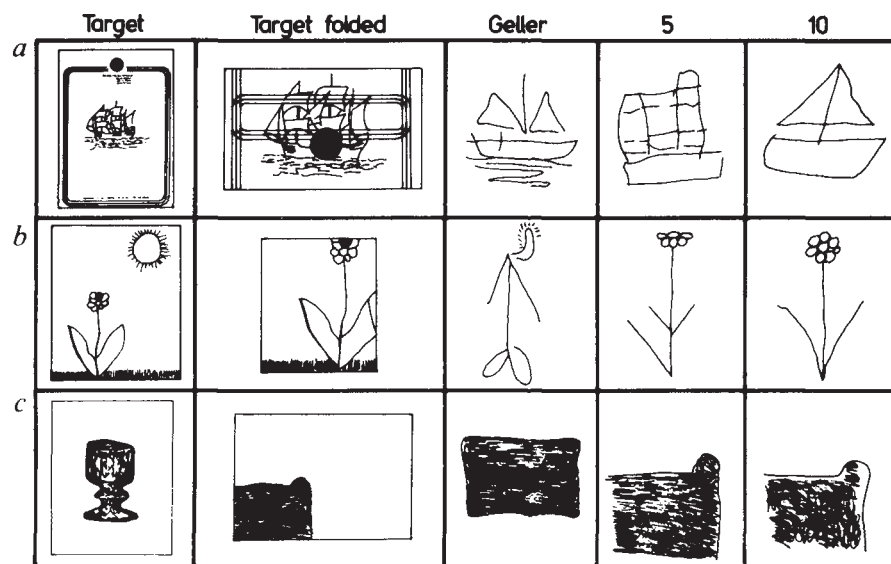


Fig. 1 Drawings presented to Geller and to non-psychic controls. The targets were inside envelopes, and Geller and the controls were allowed to see the sealed envelopes. The first and second columns show targets as they appeared in unfolded and folded states. Geller's final attempts at reproducing the drawings, which were folded and sealed inside envelopes, are shown in column 3. Columns 4 and 5 show the best results obtained from the control subjects after only 5 and 10 s of close visual inspection. The drawings were presented inside the same envelopes under similar lighting conditions. Geller, who claimed to use ESP, had taken 11.5, 85 and 18 min for *a*, *b* and *c*, respectively, but he was being observed, albeit discontinuously (see ref. 23).

they themselves believe in psi, the so-called 'sheep-goat' effect³². Hence, psi can be held to be present whatever the results, unless the belief of the investigators is itself independently controlled.

In an extreme version of the 'sheep-goat' hypothesis, some investigators have even proposed that the participants in an experiment need not be restricted to the individuals in the laboratory but could also include all of the readers of the journal which subsequently publishes the results³³. In this case the proportion of sceptics and believers among the eventual readership would determine the experimental outcome through backward causality. Similar in kind is the 'shyness effect', the tendency of metal not to bend psychically while it is being observed³⁴.

Evidence of the paranormal is held to be incompatible with materialism. Investigators throughout history have been convinced that evidence of the paranormal proves that materialism must be wrong. This was assumed by the Society of Psychical Research, one of whose early presidents, Sir William Barrett, spoke of parapsychology 'as the most valuable handmaid to religion'³⁵. J. B. Rhine³⁶, the founder of the Parapsychology Association and C. Tart³⁷, a former president, have both reiterated the religio-spiritual motive for pursuing psi research.

An immaterial 'soul' has passed out of the formal language of parapsychology, but anti-materialism is still the backbone of the underlying philosophy. A. Flew has described the profound logical difficulties with an immaterial, immortal entity which

somehow discriminates its own mental experiences from that of all others³⁸. But even putting that issue to one side, it is curious how seldom the anti-materialist assumption has been properly explained. D. E. Cooper indicated one way in which the anti-materialist argument can be constructed as a *reductio ad absurdum*, but he found this to be incoherent³⁹. In fact, it seems doubtful that materialism and ESP would be incompatible, should real evidence of ESP ever be found. As M. Scriven has pointed out, materialism can always be enlarged to absorb newly substantiated phenomena, "since the very act of substantiation demonstrates that the phenomena are indeed part of the material world, and hence that a current version of materialism must embrace them"⁴⁰.

Methodological problems

The failure of paranormal investigators to produce a single repeatable effect despite 100 years of published research is a serious matter. The hoped-for results have been described in thousands of reports, but not one can be repeated in a properly controlled replication. Yet in addition to the huge literature of unrepeatable findings, there is an inestimable number of unpublished and presumably negative results.

The most systematically investigated area is undoubtedly parapsychology. The field is professionally organized, with its own associations of accredited members and journals. Since 1969 the Parapsychological Association has been an affiliated division of the American Association for the Advancement of Science. On the sur-

face, the research sophistication of many parapsychologists seems to be as high as that of other professional researchers. The University of Edinburgh now has its own Koestler Chair of Parapsychology. Yet parapsychology is unique in that it remains permanently in search of a reliable finding. In spite of the long history of error, fraud and negative results, the practitioners remain confident that a positive result will soon be obtained. While many abortive leads have been reported in its major publications (for example, *Advances in Parapsychological Research*, *Handbook of Parapsychology*⁴¹, *Journal of Parapsychology*), there is no paradigmatic experiment in the Kuhnian sense, and every new investigator must start afresh, as though he or she is the first worker in the field.

Leading parapsychologists acknowledge the unrepeatability and admit that no single experiment has been free of error. J. Beloff⁴² and R. Morris⁴³ have concluded that the best case for psi rests on collections of experiments which, although individually flawed, reveal the undeniable presence of psi. But badly controlled experiments prove nothing, no matter how large the collection.

If any genuinely repeatable effect is ever discovered, then existing science would be modified to accommodate the new finding, which would then become an integral part of materialist science. The continued existence of 'parascience' as a separate field depends upon the investigators' creativity in searching for new, unexplained anomalies of a singularly unrepeatable kind.

How close are we to a repeatable paranormal finding? Examination of the literature suggests, not very. In systematic reviews of parapsychological experiments, C. Akers⁴⁴ and R. Hyman⁴⁵ have independently come to the same conclusion: that the research methods and evidence are too weak to establish the existence of a paranormal phenomenon.

Akers reviewed a representative sample of 54 published experiments which used unselected subjects and reported significant results. The sample included 11 experiments using the ganzfeld technique (in which the eyes and ears receive unpatterned sensory inputs), 12 hypnosis experiments, 12 studies of personality correlates, 10 studies of attitude correlates, 5 relaxation experiments and 4 meditation experiments. Akers identified seven sources of methodological error and the majority of experiments included one or more errors (see Table 1).

Hyman's analysis included 42 ganzfeld experiments reported during 1974-81. Twenty-three (55%) claimed significant evidence of psi on at least one performance measure. Giving consideration to what can be counted as an independent study, Hyman concluded that the true suc-

Table 1 Flaw analyses of representative parapsychological experiments: proportions of samples containing each flaw

Flaw	Akers' sample (ref. 44)	Hyman's sample (ref. 45)
Sensory cues	22/54	23/42
Inadequate randomization	13/54	31/42
Inappropriate statistics or multiple testing	11/54	12/42
Potential fraud	12/54	10/42
Recording errors	10/54	—
Classification or scoring errors	9/54	—
Reporting failures	10/54	25/42
One or more errors	46/54	42/42

cess rate was at most 31%. Moreover, many studies had conducted multiple statistical testing by analysing more than one performance measure and Hyman suggested that a more realistic significance level would have been as high as 0.25 instead of the nominal 0.05 level. Hence, the effective significance level and percentage of significant results are approximately equal.

Hyman's tally of procedural flaws is shown in Table 1. None of his sample was judged to be free of flaws, while Akers adjudged only eight of his sample to be flawless, but stated that none could be considered ideal.

The much-publicized experiments on remote viewing by Puthoff and Targ⁴⁶ are also invalid because of the many sensory cues⁴⁷, non-randomization⁴⁷ and inappropriate statistics⁴⁸. Tart organized a re-analysis of this research and claimed to have removed all of the sensory cues and obtained the same highly significant results⁴⁹. However, he did not in fact remove all of the cues as he had stated. Attempted replications of the remote-viewing research are either flawed or, in the case of well-controlled studies, show no evidence of ESP⁵⁰.

This review leads to only one conclusion: there is no scientific evidence of ESP. Yet millions of people throughout the world believe in the reality of ESP and other paranormal phenomena. How can these two facts be reconciled? Is science mistaken, or are folk beliefs manufactured from error and illusion?

Psychological factors

Many factors of a psychological nature foster paranormal beliefs and make them a common feature of human thinking and behaviour. Our cultural traditions are steeped in religion and magic, many features of which lend themselves to belief in supernatural agencies. Scientific thinking is a recent departure in human history and scientific ideas have had little time to affect the magical thinking from which science itself evolved.

Sociologist D. O'Keefe argues that paranormal research has evolved from within the traditions of magic which themselves evolved from religion⁵¹. The current

occult revival is seen as a reaction to the excessive rationalism which many perceive in science. O'Keefe argues that religion created the 'cloud-cuckoo land' in which magic, and thence the paranormal, can flourish. Yet scientists are often ill-prepared to provide the necessary counterbalancing rational account of the paranormal. Against this background of magico-religious entrenchment, there are some extra psychological processes that make paranormal beliefs an inevitable characteristic of human consciousness and thinking.

Mental imagery. A mental image is a quasi-perceptual experience in the absence of an objective stimulus. There are huge individual differences in the reported vividness and controllability of images. In Western cultures 1-5% of the population appears regularly to experience fantasies which seem as real as actual events even though they are entirely fictional⁵². Such individuals often experience vivid, uncontrollable 'eidetic' images of almost hallucinatory quality⁵³, are highly suggestible and can be easily hypnotized⁵². They report more putatively paranormal experiences, such as telepathy, precognition, ghosts and out-of-the-body experiences. While mental imagery has a large number of practical uses in thinking, memory and problem solving, it can also occur in altered states of consciousness in which the normal level of lucidity is no longer present⁵³.

Research conducted a century ago by E. Gurney and F. W. H. Myers described 27 cases of 'spirit communication' from deceased persons⁵⁴. Eighteen of the apparitions occurred in sleep-related states normally associated with highly vivid and autonomous images which are easily mistaken for reality. The remaining cases occurred in subjects who were fully awake and these could easily have been structural eidetic images stimulated by thought-processes of the daydreaming kind⁵³. H. Sidgwick noted that 9.9% of 17,000 subjects had experienced at least one vivid visual, auditory or tactile image of a living being or object while completely awake⁵⁵. The appearance of ghosts is shaped by cultural expectancies and beliefs about what a ghost should look like⁵⁶. Mental

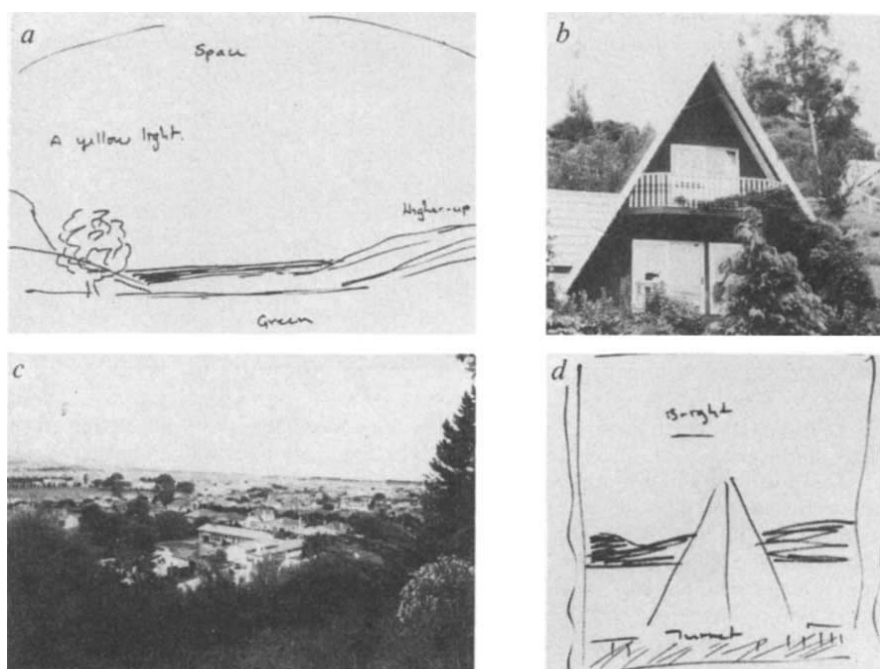


Fig. 2 *a*, A remote viewer's drawing of the target site shown in *b*, an A-frame house and steps, purportedly using ESP. When taken to the target area, the subject was delighted when he discovered the view from above the house, halfway up the steps (*c*). When an independent judge visited the target he was convinced that the correctly matching drawing was *d*. This had actually been produced for another target site, a railway station (see ref. 4).

images can be easily misinterpreted in terms of pre-existing beliefs⁵⁷.

Expectancy or mental set provides the framework within which we organize new experience. Human cognition is not a simple copying process but entails a constructive striving or 'effort after meaning'. What we experience is often more a confirmation of belief than a matter of plain fact. Beliefs are not automatically updated by the best evidence available, but have an active life of their own and fight tenaciously for their own survival. They tell us what to read, what to listen to, who to trust and how to rationalize contrary information^{4,5,57}.

Selective exposure protects beliefs from more dramatic forms of contradiction. When the mentalist U. Geller visited the city of Dunedin in New Zealand there were seven different opportunities to obtain information about his alleged psychic abilities: four media interviews, two newspaper stories and one stage performance. Of 17 subjects who, before Geller's visit, were already 'believers', 15 selected three or more of the available exposures. Of 20 'non-believers', only 10 selected as many as three exposures ($\chi^2(1) = 6.13$; $P < 0.02$).

A further problem is that when we are exposed to relevant information, our opinion revisions are often less than optimal, and we act like conservative Bayesians⁵⁸, with a confirmation bias⁵⁹. In a recent 'ESP' demonstration to a class of 226 psychology students, presented as an exercise in observation, I performed five

mentalists' tricks consisting of: (1) correctly naming a colour written out of sight; (2) correctly transmitting a colour name to a volunteer who, like me, had not previously seen it; (3) helping a volunteer correctly to read messages sealed inside envelopes or to appear to transmit messages to me; (4) producing bent keys which I had not previously touched; and (5) moving or stopping the hands of a watch in a mysterious manner.

Although at no time did I claim to be psychic, 90% of the class stated that I had demonstrated psychic ability. When the results from subjects who had previously been classified as 'believers' and 'sceptics' were analysed separately, 79% of believers thought at least three of the five effects were psychic compared with only 43% of sceptics ($P < 0.001$).

Naturally, we often encounter information that is unexpected or ambiguous. In such instances, there is a second line of defence: the data can be selectively perceived or even misperceived so that they still appear to support our beliefs by 'subjective validation'⁴. One illustration of this powerful cognitive defence in the context of ESP research is the strong conviction that one has successfully viewed a complex target site by ESP in a remote-viewing experiment even when one is completely wrong (Fig. 2).

There are many now-classic examples of subjective validation: the prophecies of the Delphic Oracle and Nostradamus⁶⁰, the discovery of N-rays⁶¹, phlogiston, Vulcan, the canals on Mars, flying saucers,

Freud's interpretation of dreams, prejudice, faith-healing, the placebo effect, bone pointing and the 'evil eye'. Beliefs of all kinds tend to be self-perpetuating. **Coincidences.** Psi phenomena consist of an experience, image or thought matched by some other similar experience, image or thought. Collections of such coincidences have been published by A. Koestler⁶², L. Rhine,⁶³ and others based on the assumption that odd-matches of events cannot occur purely by chance.

Probability theory shows that an event which is improbable over a short run can become highly probable over the long run. If five coins are tossed all at once on a single occasion the probability of obtaining five heads is 2^{-5} or approximately 0.03. If the coin tossing is repeated 100 times the probability of five heads somewhere in the series is approximately 0.96.

The principle of the long run is easy to grasp in simple situations but much less visible in the more chaotic world of spontaneous human experience. Calculation shows how easily Koestler could obtain his 40-plus odd-match anecdotes. Assuming that in an ordinary day a person can recall 100 distinct events, there are $^{100}C_2$ or 4,950 pairs of events per day. Odd-matches can be remembered for years, perhaps 10 yr or 3,650 days. If Koestler knew 1,000 people, he could draw upon a total pool of $4,950 \times 3,650 \times 1,000$, or more than 18×10^9 pairs of events. That Koestler obtained 40 striking odd-matches seems hardly surprising.

Koestler's fallacy (see ref. 4) is certainly not unique to him, although he was one of a small group of analysts who wanted to make a scientific revolution out of it. The fallacy is widespread and several biases contribute to it. First, we notice and remember odd-matches. Second, we do not notice non-matches. This triggers the short-run illusion that makes the odd-match seem improbable. Third, we are normally poor estimators of probabilities, especially for combinations of events.

Unseen causes. Another class of psychic-looking experiences is generated by invisible chains of cause and effect which bias the probabilities away from chance levels. Failure to randomize target stimuli properly in ESP experiments is a good example of this. Thus, Tart reported a successful ESP experiment in which his subjects learned to score above chance in guessing which of 10 digits was displayed by an apparatus in another room following the presentation of feedback⁴. The random number generator mistakenly avoided using the same digit twice in succession, a bias which is matched by the pervasive 'gambler's fallacy'. When Tart removed this bias, the 'ESP' also disappeared⁶⁵.

Another unseen factor, used by illusionists, is the 'population stereotype'. The performer 'sends a message' to the

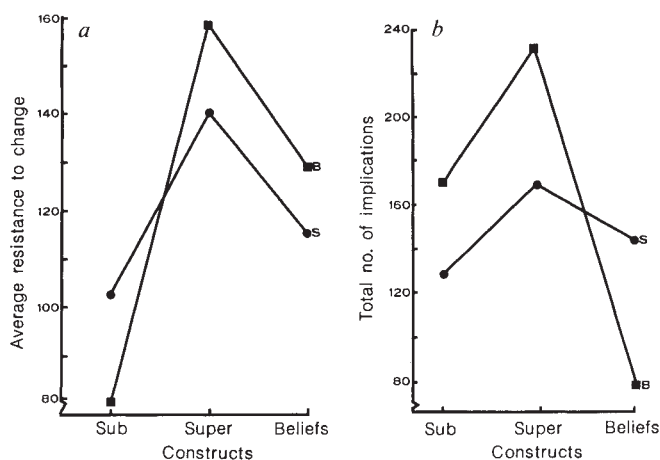


Fig. 3 *a*, The average resistance to change scores for 'believers' (B) and 'sceptics' (S) for subordinate (Sub), superordinate (Super) and belief constructs. Believers showed a significantly higher resistance to change for superordinate constructs than sceptics. *b*, The total number of implications for the same sets of constructs. The interaction between group and construct-type was highly significant ($F(2, 42) = 9.14$, $P < 0.005$). Believers saw significantly more implications in changing positions on superordinate constructs than did the sceptics, while sceptics saw significantly more implications in changing their beliefs than believers ($P < 0.05$ for both comparisons).

audience, saying "I am thinking of a number between 1 and 50, both digits are odd, and different". Controlled experiments show that the most common response for the 1-50 problem is 37, which accounts for 30-35% of all responses, and the second most common response is 35 (20-25%)⁴. If the performer always says he had been thinking of 35 and then changed his mind to 37, at least 50% of the audience will be thinking of the 'correct' number.

Human beings never behave randomly. Our experiences contain many culturally shared elements such that particular items are associated with particular verbal contexts. This causes associative networks to be set up and a tendency towards non-random, stereotypical responses even when there is freedom to choose.

Other unnoticed causes of putatively psychic effects include subliminal and non-verbal sensory cues⁶⁶ which may lead to common thought patterns in different people, presenting the illusion of telepathy.

The 'will to believe'. What factors differentiate believer from sceptic? Psychologists down the ages have puzzled over the question of what motivates different world-views and the so-called will to believe. Research conducted by J. Waugh used Kelly's personal construct theory. In this framework⁶⁷, people vary in the quality and extent of their investigatory procedures so that, while some may be working to establish an ordered and meaningful world which is not highly predictable or readily explained, others may be content that they already have all the necessary explanatory constructs.

In Kelly's theory, each individual deals with the world in terms of a hierarchical

system of constructs with which people, objects and events are compared, contrasted and predicted. Core constructs have relatively superordinate positions and a large range of convenience while peripheral constructs are relatively subordinate and more easily altered. Waugh compared the personal construct systems of sceptics and believers in the paranormal using a belief questionnaire. Ten subordinate and ten superordinate constructs were generated using standard procedures and each subject's constructs were tested for their relative resistance to change and the number of implications entailed by changing the subject's preferred pole on the 20 constructs and 10 paranormal beliefs (Fig. 3).

Believers' core constructs were significantly more resistant to change and there was a parallel difference in the number of implications resulting from changes at the superordinate level. Compared with sceptics, believers seem to possess much tighter construct systems in which any change at the core level implies a significantly greater upheaval or threat. Waugh also found that believers had significantly higher neuroticism scores than sceptics (see also ref. 68). These data are congruent with those reported by Zusne and Jones⁵⁷ who found that believers are less flexible than sceptics when confronted with disconfirming evidence. Content analyses of believers' construct systems indicate the presence of spiritual, non-materialist constructs at superordinate level. Such core constructs are not easily shaken because they are closed off from empirical considerations and appear to be impermeable to rational persuasion. Hence the feeling of futility experienced in trying to hold rational discussion

between believer and sceptic; one could well be arguing about the existence of God. Belief in the paranormal is metaphysical and therefore not subject to the constraints of empirically based science.

Parascience has all the qualities of a magical system while wearing the mantle of science. Until any significant discoveries are made, science can justifiably ignore it, but it is important to say why: parascience is a pseudo-scientific system of untestable beliefs steeped in illusion, error and fraud.

I thank Jerry Andrus, Bob Audley, Ray Hyman, A. R. Jonckheere, Peter McKellar, J. Randi, Christopher Scott, Jean Waugh and many colleagues in CSICOP for useful discussions and information. The late Richard Kammann contributed substantially in the earlier stages of this research.

David F. Marks is at the Department of Psychology, University of Otago, PO Box 56, Dunedin, New Zealand.

1. Kurtz, P. *Skeptical Inquirer* 3, 14-32 (1978).
2. Diaconis, P. *Science* 201, 131-136 (1978).
3. Hansel, C. E. M. *ESP and Parapsychology: A Critical Reevaluation* (Prometheus, Buffalo, 1980).
4. Marks, D. & Kammann, R. *The Psychology of the Psychic* (Prometheus, Buffalo, 1980).
5. Alcock, J. E. *Parapsychology: Science of Magic?* (Pergamon, Oxford, 1981).
6. Frazier, K. (ed.). *Paranormal Borderlands of Science* (Prometheus, Buffalo, 1981).
7. Gardner, M. *Science Good Bad and Bogus* (Prometheus, Buffalo, 1981).
8. Randi, J. *Flim-Flam! Psychics, ESP, Unicorns, and Other Delusions* (Prometheus, Buffalo, 1982).
9. Kurtz, P. *Skeptical Inquirer* 8, 239-246 (1984).
10. Frazier, K. (ed.). *Science Confronts the Paranormal* (Prometheus, Buffalo, 1985).
11. Kurtz, P. (ed.). *A Skeptics Handbook of Parapsychology* (Prometheus, Buffalo, 1985).
12. Bunge, M. *Skeptical Inquirer* 9, 36-46 (1984).
13. Tiller, W. A. *New Scientist*, 62, 160-163 (1974).
14. Pehk, J. O., Kyler, H. J. & Faust, D. L. *Science* 194, 263-270 (1976).
15. Houdini *Miracle Mongers and Their Methods* (Prometheus, Buffalo, 1981).
16. Leikind, B. J. & McCarthy, W. J. *Skeptical Inquirer* 10, 23-34 (1985).
17. Vogt, E. H. & Hyman, R. *Waterwitching USA* 2nd edn (Chicago University Press, 1979).
18. Randi, J. *Skeptical Inquirer* 8, 329-333 (1984).
19. Martin, M. *Skeptical Inquirer* 8, 138-140 (1983).
20. Randi, J. *The Magic of Uri Geller* (Ballantine, New York, 1975).
21. Fuller, U. *Confessions of a Psychic* (Karl Fulves, Box 433, Teaneck, New Jersey, 1975).
22. Fuller, U. *Further Confessions of a Psychic* (Karl Fulves, New Jersey, 1980).
23. Marks, D. & Kammann, R. *Zetetic* 1(2), 9-17 (1977).
24. Hyman, R. *Zetetic* 1(2), 18-37 (1977).
25. Soal, S. G. & Goldney, K. M. *Proc. Soc. psychical Res.* 47, 21-150 (1943).
26. Scott, C. & Haskell, P. *Nature* 245, 52-54 (1974).
27. Marwick, B. *Proc. Soc. psychical Res.* 56, 250-281 (1978).
28. Hoebens, P. H. *Skeptical Inquirer* 6, 32-40 (1981).
29. De Mille, R. *Castaneda's Journey* 2nd ed (Capra, Santa Barbara, 1978).
30. Skinner, B. F. *Science and Human Behavior* (Macmillan, New York, 1953).
31. Bunge, M. *Method, Model and Matter* (Reidel, Dordrecht, 1973).
32. Schneider, G. R. & McConnell, R. A. *ESP and Personality Patterns* (Yale University Press, 1958).
33. Collins, H. M. & Pinch, T. J. *Frames of Meaning: The Social Construction of Extraordinary Science* (Routledge & Kegan Paul, London, 1982).
34. Taylor, J. *Superminds: An Inquiry into the Paranormal* (Macmillan, London, 1975).
35. Barratt, W. *Proc. Soc. psychical Res.* 34, 275-297 (1924).
36. Rhine, J. B. *The Reach of Mind*, 209-214 (Sloane, New York, 1947).
37. Tart, C. T. *Psi: Scientific Studies of the Psychic Realm*, vii-viii (Dutton, New York, 1977).
38. Flew, A. in *Science, Pseudo-Science and Society* (eds Hanen,

- M. P., Osler, M. J. & Weyant, R. G.) 55-75 (Wilfrid Laurier University Press, Waterloo, 1980).
39. Cooper, D. E. in *Philosophy and Psychological Research* (ed. Thakur, S. C.) 59-80 (Allen & Unwin, London, 1976).
40. Scriven, M. in *Philosophy and Psychological Research* (ed. Thakur, S. C.) 181-194 (Allen & Unwin, London, 1976).
41. Wolman, B. J. (ed.) *Handbook of Parapsychology* (Van Nostrand, New York, 1977).
42. Beloff, J. *Zetetic Scholar* 6, 90-94 (1980).
43. Morris, R. L. *J. Am. Soc. psychical Res.* 74, 425-443 (1980).
44. Akers, C. in *Advances in Parapsychological Research* Vol. 4 (ed. Krippner, S.) 112-164 (McFarland, Jefferson, North Carolina, 1984).
45. Hyman, R. *J. Parapsychol.* 49, 3-49 (1985).
46. Targ, R. & Puthoff, H. E. *Nature* 252, 602-607 (1974).
47. Marks, D. F. & Kammann, R. *Nature* 274, 680-681 (1978).
48. Hyman, R. *Skeptical Inquirer* 9, 125-145 (1984-5).
49. Tart, C. T., Puthoff, H. E. & Targ, R. *Nature* 284, 191 (1980).
50. Marks, D. F. *Skeptical Inquirer* 6, 18-29 (1982).
51. O'Keefe, D. *Stolen Lightning* (Robertson, Oxford, 1982).
52. Wilson, S. C. & Barber, T. X. in *Imagery: Current Theory, Research and Application* (ed. Sheikh, A. A.) 340-387 (Wiley, New York, 1983).
53. Marks, D. & McKellar, P. *J. mental imagery* 6, 1-124 (1982).
54. Gurney, E. & Myers, F. W. H. *Proc. Soc. psychical Res.* 5, 403-485 (1889).
55. Sidgwick, H. *Proc. Soc. psychical Res.* 10, 25-422 (1894).
56. Finucane, R. C. *Appearances of the Dead* (Prometheus, Buffalo, 1985).
57. Zusne, L. & Jones, W. H. *Anomalous Psychology* (Erlbaum, Hillsdale, New Jersey, 1982).
58. Edwards, W. in *Formal Representation of Human Judgement* (ed. Kleinmuntz, B.) 17-52 (Wiley, New York, 1968).
59. Nisbett, R. & Ross, L. *Human Inference: Strategies and Shortcomings of Social Judgment* (Prentice-Hall, Englewood Cliffs, 1980).
60. Hoebens, P. H. *Skeptical Inquirer* 1, 38-45 (1982).
61. Klass, P. J. *Zetetic* 2, 57-61 (1977).
62. Koestler, A. *The Roots of Coincidence* (Hutchinson, London, 1972).
63. Rhine, L. E. *J. Parapsychol.* 15, 164-190 (1951).
64. Tart, C. T. *Learning to Use Extrasensory Perception* (University of Chicago Press, 1976).
65. Tart, C. T., Palmer, J. & Redington, D. J. *J. Am. Soc. psychical Res.* 73, 151-165 (1978).
66. Dixon, N. F. *Preconscious Processing* (Wiley, Chichester, 1981).
67. Fransella, F. & Bannister, D. *A Manual for Repertory Grid Technique* (Academic, London, 1977).
68. Windholz, G. & Diamant, L. *Bull. psychon. Sci.* 3, 125-126 (1974).

REVIEW ARTICLE

Heavy-electron metals

Z. Fisk, H. R. Ott*, T. M. Rice* & J. L. Smith

Center for Materials Science, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

A new class of metals has been found in which the electrons have effective masses orders of magnitude larger than the free-electron mass. Some of these metals are superconducting at low temperatures. This superconductivity seems to be unconventional, with an underlying mechanism different from that in all other known superconductors.

AMONG the main developments in physics since the 1930s have been the discovery and exploration of new ground states of condensed matter. Each of these discoveries opened a new chapter in the physics of condensed matter; recent examples are spin- and charge-density waves in metals, the superfluid state of liquid ^3He and the quantized Hall effect. Very recently an exciting new class of metallic materials has been discovered with remarkable properties and showing signs of further new ground states. (A general compilation of data, with references current as of mid-1984, can be found in ref. 1.)

The terms 'heavy-electron metal' and 'heavy-fermion system' have been introduced to describe materials in which the electronic states have a characteristic energy orders of magnitude smaller than in ordinary metals. If we write the energy $\epsilon(k)$ in a free-electron form ($\epsilon(k) = \hbar^2 k^2 / 2m^*$), then since the wave-vectors k of the electron determined by the interatomic spacing are not much different, the effective mass m^* must be orders of magnitude larger than the free-electron value and in some cases m^* is a fair fraction of the proton mass. These materials are intermetallic compounds in which one of the constituents is a rare-earth or actinide atom, with partially filled 4f- or 5f-electron shells. At high temperatures these materials behave as if the f-electrons were localized on their atomic sites, as in conventional rare-earth and actinide compounds, where any itinerant electrons are in states derived from loosely bound atomic s-, p- and d-orbitals. As the conventional materials are cooled, the atomic moments due to the f-electrons order spontaneously, mostly antiferromagnetically, less often ferromagnetically. By contrast, in the heavy-electron systems some of the f-electrons become itinerant at low temperatures and form a metallic state with the characteristics described above.

Recently the exciting discovery was made that in some of these new materials the heavy electrons form a superconducting state at very low temperatures^{3,4,6}. Superconductivity in ordinary metals is associated with an instability of itinerant electrons; thus it was surprising that it should also occur where normal-

state properties are dominated by nearly localized electrons. Various features of this superconducting state are unusual, leading theorists to speculate that not only is the mechanism driving this superconducting transition unconventional, but also that the configuration of the superconducting state is different from that of an ordinary superconductor. Instead of an interaction between electrons that is mediated by phonons (lattice-vibrational quanta), which leads to an essentially isotropic gap in the spectrum of electronic excitations, it is envisaged that a Coulomb interaction between heavy electrons induces a superconducting state in which the energy gap is strongly anisotropic. If this could be definitely established, it would be the culmination of a long search for this phenomenon.

There is no doubt that a proper understanding of this superconducting state requires a clear understanding of the preceding normal state with its remarkable properties. Early theories of the properties of metals always assumed that, whereas the conduction electrons interact with the ionic lattice forming the solid, they do not interact at all among themselves. Quantum statistics then determine the low-temperature properties of this electron gas, two of which are of particular importance in the context of our discussion. The specific heat of this electron gas c_p varies linearly with temperature as $T \rightarrow 0$ K (that is, $c_p = \gamma T$). The low-temperature magnetic susceptibility, χ , is independent of temperature. In this simple theory the ratio

$$\chi / \gamma = 3 \mu_B^2 / \pi^2 k_B^2 \quad (1)$$

is obviously a universal number. $\mu_B = 9.27 \times 10^{-21}$ erg G^{-1} is the Bohr magneton and $k_B = 1.38 \times 10^{-16}$ erg K^{-1} is the Boltzmann constant. The factor which determines the magnitude of both χ and γ is the density of electronic states per unit energy, $N(E_F)$, at the Fermi energy E_F (E_F is the energy up to which all possible states of the electron gas are occupied at $T = 0$ K.) Hence $N(E_F)$ varies inversely with the characteristic energy of the electrons, leading $N(E_F)$ to be proportional to m^* . This simple concept is quite adequate to describe the qualitative low-temperature features of simple metals, for which γ is of the order of 1 mJ $\text{mol}^{-1} \text{K}^{-1}$, $\chi \approx 10^{-5}$ e.m.u. mol^{-1} and $T_F = E_F / k_B$, the Fermi temperature, is $\sim 10^4$ – 10^5 K. The experimental facts quoted in

* Permanent address: Eidg. Tech. Hochschule, ETH-Hönggerberg, CH-8093 Zürich, Switzerland.

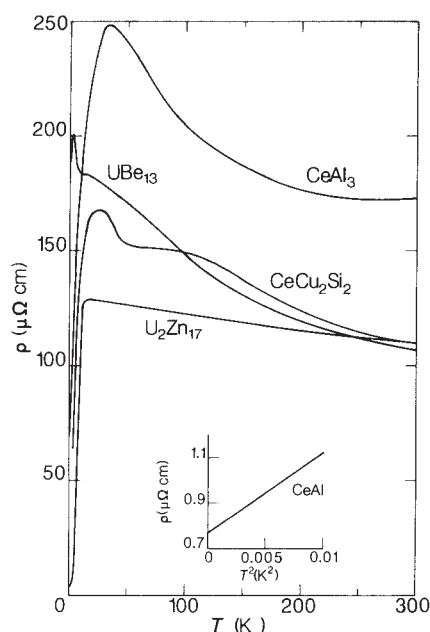


Fig. 1 Temperature dependence of the electrical resistivity ρ of four different heavy-electron compounds below room temperature. The high values indicate very strong scattering of the electrons but the distinct features and the resistivity decrease at low temperatures demonstrate that these are not simply 'dirty' metals. The inset reveals the T^2 dependence of the ρ of CeAl₃ at very low temperatures.

the next section should not only make it clear that these parameters are very different in this new class of materials but also demonstrate the intriguing transition from localized to itinerant behaviour of these heavy electrons.

Normal-state properties

First we consider the electrical resistivity ρ . In ordinary metals ρ decreases rapidly with decreasing temperature below 300 K, from values that are ~ 1 – $10 \mu\Omega \text{ cm}$. In heavy-electron materials $\rho \approx 100 \mu\Omega \text{ cm}$ at room temperature. It often increases with decreasing temperature and only after passing over a maximum at $T \leq 50 \text{ K}$ does ρ decrease to low values as $T \rightarrow 0 \text{ K}$. Figure 1 shows examples of this behaviour in CeAl₃, CeCu₂Si₂, UBe₁₃ and U₂Zn₁₇. CeAl₃ was the first to be identified as a heavy-electron metal². It stays normal to the lowest temperature investigated ($\sim 10^{-2} \text{ K}$) and therefore can be used to study the heavy electrons in their normal state. For CeAl₃ it is found that although ρ reaches a maximum value of the order of $200 \mu\Omega \text{ cm}$ at 35 K, it drops to less than $1 \mu\Omega \text{ cm}$ at $T = 0 \text{ K}$, where this residual value of the resistivity is determined by impurities and imperfections of the crystal lattice as in ordinary metals. As an intriguing fact it was noted that $\rho \propto T^2$ at the lowest temperatures with a remarkably large pre-factor of $35 \mu\Omega \text{ cm K}^{-2}$, indicating a very effective, temperature-dependent scattering mechanism even at these very low temperatures. (See inset, Fig. 1.)

The first observation of superconductivity in heavy-electron metals was made in CeCu₂Si₂ (ref. 3), and this was quite unexpected. As shown in Fig. 1, its $\rho(T)$ is very anomalous in the normal state. The increase of ρ seems to saturate below 100 K but reaches a pronounced maximum in the vicinity of 10 K before it decreases rapidly and finally vanishes discontinuously at the superconducting transition around 0.5 K. UBe₁₃ is a 5f-electron-based heavy-electron system, and is also becomes superconducting below 1 K (ref. 4). Its $\rho(T)$ is indeed very much like that of CeCu₂Si₂, but the characteristic structures (shoulder and peak) are shifted to lower temperatures. The decrease of ρ below 2.5 K is interrupted by the superconducting transition at

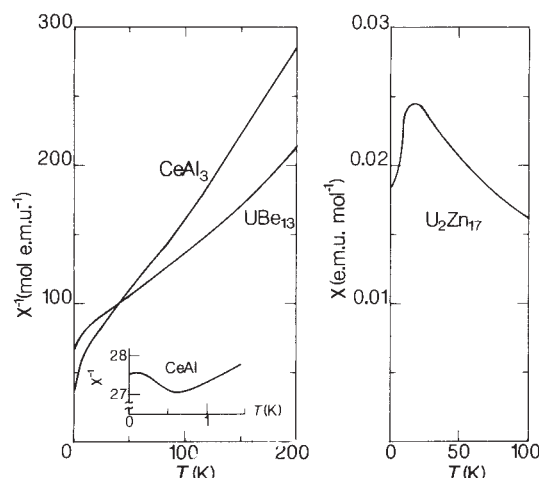


Fig. 2 The Curie-Weiss behaviour of the magnetic susceptibilities χ of CeAl₃ and UBe₁₃ above 100 K is evidence for localized f-electron moments. The low-temperature saturation (inset) for CeAl₃, however, demonstrates the itineracy of the f-electrons below 1.5 K. For U₂Zn₁₇, the tendency to saturation is interrupted by a sharp drop of χ due to antiferromagnetic ordering.

0.9 K. Finally we consider the $\rho(T)$ of U₂Zn₁₇, a substance that orders magnetically from a heavy-electron state⁵. Here the maximum resistivity is observed at 17 K and the rapid decrease of almost two orders of magnitude occurs below 10 K.

Magnetic moments that are due to partially filled 5f-electron shells of individual atoms give rise to the familiar Curie-Weiss-type behaviour of the magnetic susceptibility χ with varying temperature; that is, $\chi^{-1} \propto (T - \theta_p)$, where θ_p is the paramagnetic Curie-Weiss temperature. This type of behaviour is clearly reflected in the plots of χ^{-1} versus T for CeAl₃ and UBe₁₃ in Fig. 2. Deviations from this simple behaviour are apparent at $T < 100 \text{ K}$ but they can still be ascribed to local f-electron configurations. In both materials χ saturates only at $T \leq 1.5 \text{ K}$ (see inset, Fig. 2). In UBe₁₃, χ jumps discontinuously to the ideal diamagnetic value of $(4\pi)^{-1}$ at the superconducting transition (at $T_c \sim 0.9 \text{ K}$). For U₂Zn₁₇, we show $\chi(T)$ below 100 K; note the trend to saturation and the broad maximum at 17 K. The sharp decrease just below 10 K is due to an antiferromagnetic phase transition. The saturation values of χ for all three materials are very large compared with those of normal metals ($3.6 \times 10^{-2} \text{ emu mol}^{-1}$ CeAl₃; $1.5 \times 10^{-2} \text{ emu mol}^{-1}$ UBe₁₃; $2.45 \times 10^{-2} \text{ emu mol}^{-1}$ U₂Zn₁₇).

As mentioned above, the electronic specific heat of ordinary metals can be described by a single term, $c_p^{\text{el}} = \gamma T$ at temperatures below 10 K. That this is not the case for heavy-electron materials is shown in Fig. 3, where we plot c_p/T against T for CeAl₃, CeCu₂Si₂ and UBe₁₃ for temperatures between 1 and 10 K (Fig. 3a) and for UPt₃ (ref. 6) between 1 and 17 K (Fig. 3b). For ordinary metals, such a plot would result in a straight line with a positive slope and an ordinate $\approx 1 \text{ mJ mol}^{-1} \text{ K}^{-2}$ at $T = 0 \text{ K}$. For heavy-electron materials in general, the ratio c_p/T rises considerably with decreasing temperature in this temperature range, starting from values of about $100 \text{ mJ mol}^{-1} \text{ K}^{-2}$. The temperature dependence itself might be loosely interpreted by stating that the γ parameter is no longer constant but is strongly temperature-dependent. Below 1 K, the c_p/T of CeAl₃ rises to a maximum of more than $2 \text{ J mol}^{-1} \text{ K}^{-2}$ at about 0.5 K and finally saturates at $\sim 1.6 \text{ J mol}^{-1} \text{ K}^{-2}$ as T approaches 0 K. For both CeCu₂Si₂ and UBe₁₃, c_p/T rises in a very similar way but then saturates around 1 K and remains almost constant until a specific-heat anomaly indicates the transition to the superconducting state, as is shown, for example, for UBe₁₃ in Fig. 3c.

As outlined in the introduction, these very large values of the c_p^{el}/T ratio (or γ value) as T approaches 0 K must be

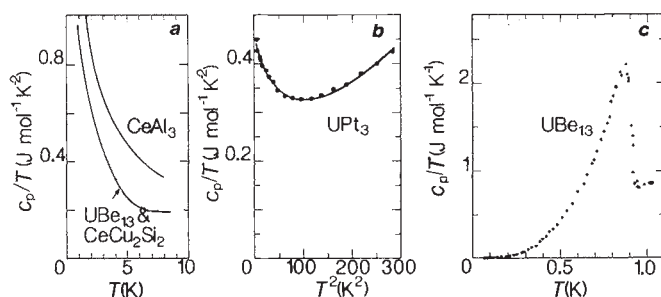


Fig. 3 The low-temperature specific heat c_p of these materials is unusually large. The values of $T \approx 1$ K are more than a hundred times larger than for ordinary metals and originate in the strong interactions between electrons. An analytic form describing the experimental data has been found only for UPt_3 ; this is indicated by the solid line in *b*. The specific-heat anomaly of UBe_{13} (*c*) below 1 K is due to the transition to the superconducting state.

interpreted as evidence for enormous densities of states at E_F in the electronic spectrum of these substances. This in turn implies very large effective masses for these electrons, and hence provides the name now generally used for these systems. Although both γ and χ are enhanced in these materials above the values seen in ordinary metals, their ratio χ/γ is comparable to that defined in equation (1).

Phase transitions within heavy-electron state

Superconductivity. The discovery of superconductivity in $CeCu_2Si_2$ by Steglich *et al.*³ was surprising because it violated many long-held beliefs in the field. In particular, the Curie-Weiss-type high-temperature magnetic susceptibility suggested local magnetic moments on the Ce atoms, and these are known to depress conventional s-state (or singlet) superconductivity by breaking up the Cooper electron pairs.

The specific-heat anomaly (shown for UBe_{13} in Fig. 3*c*) at the superconducting transition temperature (critical temperature, T_c scales with the normal-state value, giving not only convincing proof of bulk superconductivity but also hard evidence that $\gamma (= c_p/T)$ is really due to a large density of states of itinerant electrons. The superconducting energy gap opens in the large-mass band; the f-electrons are therefore intimately involved in the superconductivity. It is important to distinguish this behaviour from that of the Chevrel-phase and ternary-boride magnetic superconductors⁷. In these latter we have an independent sublattice of magnetic ions, the magnetic order parameter of which competes with the superconducting order parameter of the essentially separate conduction electron system; in the heavy-electron materials it is one and the same set of electrons which deliberates between superconductivity and magnetism.

Arguments given below suggest the possibility that the superconductivity observed in these materials may not be of the s-state, isotropic type usual in the Bardeen-Cooper-Schrieffer (BCS) theory of superconductivity⁸. It may instead be an anisotropic p-, or possibly d-state type also possible in a generalized BCS theory, and now quite firmly established as the pairing state in superfluid 3He (ref. 9) but previously unknown in metals. Anisotropic superconducting states may have zeroes of the superconducting energy gap on the Fermi surface and these lead to non-exponential temperature dependences below T_c in experimental measurements such as specific heat, ultrasonic attenuation, thermal conductivity and NMR relaxation rates.

The experimental situation can be illustrated with data for UBe_{13} and UPt_3 . In UBe_{13} , the superconducting specific-heat jump at T_c (see Fig. 3*c*) is $\sim 50\%$ larger than the weak-coupling BCS prediction, placing it very squarely in the strong-coupling regime. Below T_c , the specific heat varies with a power law close to T^3 (ref. 10). Ultrasonic attenuation¹¹ and NMR relaxation rates¹² also follow power laws, as shown in Fig. 4. Deviations

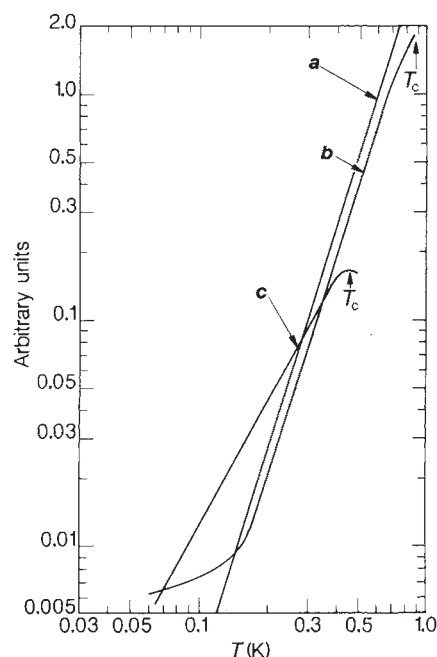


Fig. 4 Examples of the power-law-type temperature dependence of various properties in the superconducting state of UBe_{13} and UPt_3 , indicating anisotropic gaps in the electronic excitation spectrum. Conventional isotropic superconductivity would result in exponential temperature dependences for all these physical quantities. *a*, Specific heat of UBe_{13} ($\propto T^3$); *b*, inverse spin-lattice relaxation rate of 9Be in UBe_{13} ($\propto T^3$); *c*, ultrasonic attenuation in UPt_3 ($\propto T^2$).

from exponential behaviour are found in conventional superconductors containing many magnetic impurities, but the behaviour of the heavy-electron superconductors cannot be explained in this way. In addition, ultrasonic attenuation experiments on UBe_{13} (ref. 13) reveal a peak in attenuation just below T_c , a feature not observed previously in any other superconductor, again suggesting an unconventional form of superconductivity.

The effect of substitutional impurities on the T_c of superconductors is often informative, because impurities carrying local magnetic moments depress T_c in conventional superconductors in a way which varies from impurity to impurity in a predictable way. Local moments are, as mentioned above, strongly pair-breaking for s-state superconductors. In UBe_{13} , the substitution for U of rare-earth impurities with local moments leads to fairly rapid depressions of T_c , but there is no observable difference between those rare-earths which carry local moments and those which do not. Lutetium, for example, at a concentration of 3 atom per cent (at. %) in UBe_{13} has pushed T_c to below 20 mK.

Thorium substitution for U is quite peculiar (Fig. 5). Instead of a monotonic depression of T_c as observed for the rare-earth substitutions in UBe_{13} , an initial depression is followed by a plateau or even a slight rise in T_c at 0.6 K between 2 and 4 at. % Th, after which T_c again decreases, the specific-heat jump being sizeably reduced by the time the Th concentration reaches 6 at. %. Completely unexpected was the finding that there are actually two consecutive second-order phase transitions of comparable magnitude in the plateau region, the lower-temperature one being at approximately 0.4 K (ref. 15). Experiment shows that these samples remain superconducting below the second transition.

This second transition involves an ultrasonic-attenuation anomaly that is two orders of magnitude larger than the first, superconducting one¹⁶. The similarity to anomalies observed by this technique at magnetic transitions led these authors to suggest that it might be a spin-density-wave transition. The ordered

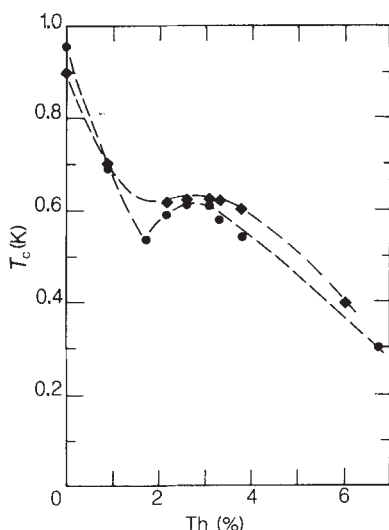


Fig. 5 Influence of small amounts of Th substitution for U in UBe_{13} (ref. 15). Non-magnetic impurities in superconductors usually lead to smooth depressions of the critical temperature T_c . For $0.02 < x_{\text{Th}} < 0.04$, a second phase transition occurs at about 0.4 K. Its nature is still not well established. ●, T_c , obtained from magnetic susceptibility; ◆, T_c , obtained from specific heat.

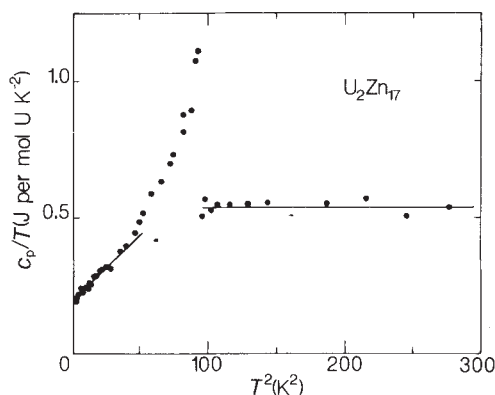


Fig. 6 Electronic specific heat of U_2Zn_{17} at low temperatures. The anomaly is due to the antiferromagnetic ordering. The finite ordinate at $T = 0$ K indicates that 40% of the large electronic density of states is unaffected by the transition.

magnetic moment would be $\sim 0.01 \mu_B/\text{U}$, consistent with the upper limit set by both NMR and neutron scattering experiments. However, there appears to be no theory at present which can explain how a spin-density wave and superconductivity could co-exist in the same large-mass band.

Another possibility is that the second phase transition is to a second superconducting state. This certainly would require some kind of unconventional superconductivity. The remaining possibility of a simple structural phase transition is unlikely because, by experience, impurities tend to suppress, not enhance, such transitions.

Magnetism. We illustrate the occurrence of magnetic order in heavy-electron metals with the example of U_2Zn_{17} . It orders at a transition temperature $T_N = 9.8$ K into a simple antiferromagnetic structure with an ordered moment of approximately $0.8 \mu_B$ (ref. 17). This value is considerably smaller than the Curie-Weiss effective moment of $3.3 \mu_B$. Just above T_N , $\gamma = 504$ mJ per mol U K^{-2} , falling to a limiting value of 198 mJ per mol U K^{-2} as $T \rightarrow 0$ K as shown in Fig. 6.

The specific-heat anomaly at T_N shows that the magnetism

involves the heavy electrons. Also note that the residual γ is still very large, indicating that some heavy electrons survive in the magnetically ordered state. The magnetic transition can be thought of as removing 60% of the Fermi surface, and this seems to be a general feature of the other known magnetically ordering heavy-electron systems. These systems resemble itinerant-electron magnets such as chromium, not local-moment magnets. This resemblance is strengthened by the fact that the magnetic order in U_2Zn_{17} is strongly affected by substitution on the Zn sites, completely unlike usual local-moment behaviour and very reminiscent of what is observed in Cr.

Theoretical understanding

The fascinating and unusual properties of the heavy-electron metals have attracted the attention of many theorists. Their work can be divided into two broad categories: first, attempts to construct a microscopic theory of the normal-state properties, and second, the use of phenomenological models to examine the experimental consequences of, and the tests for, unconventional superconductivity. Several summaries of current theories have recently been published¹⁸⁻²⁰.

Normal state. The starting point for the microscopic theories is a periodic version of the model hamiltonian introduced many years ago by P. W. Anderson²¹, which serves to describe a single rare-earth or actinide impurity in a conventional metal. There are three parts to this model. One part describes the conduction-electron states derived from s- and p-states as free-electron states and all effects of the Coulomb interaction among these electrons are ignored. In the second part the f-atomic states are described as a set of degenerate energy levels, but because these f-states are spatially localized in the vicinity of the atomic core, the Coulomb interaction between electrons in these states is large and must be included. Consider what will happen if the energy of the degenerate set of f-levels is considerably below the Fermi energy E_F , but the strong Coulomb interaction prevents the f-states from being completely filled. There are now several ground states of the system, because the partially filled f-level can have more than one configuration. The third and complicating part of the Anderson model is the term which describes so-called hybridization, or electron transfer between the f-states and the conduction states. This hybridization is in general a small term, but it allows the possibility of making transitions within this degenerate (in energy) set of ground states so that one particular combination becomes the true ground state. This process has been extensively studied for the single-site impurity case, where it is known as the Kondo effect, after J. Kondo, who posed the problem more than twenty years ago²². After many years of theoretical work, exact analytic solutions for this case have been obtained. The solution is characterized by a non-degenerate ground state in which the magnetic susceptibility χ is finite and the entropy associated with the partially filled f-states vanishes linearly in temperature so that the excess specific heat (associated with the impurity) is $\sim \gamma T$. The ratio χ/γ can be enhanced by up to a factor of 2 from the universal form in equation (1). The mixing between the conduction states and the f-states leads to a very large electrical resistance in the limit as $T \rightarrow 0$ K.

In the heavy-electron metals the f-state atoms sit on periodic lattice sites and the problem is no longer exactly soluble. Some authors (see refs 18-20) have suggested simply applying what is known about the single-site impurity problem by considering a set of single-site impurities on the lattice. The lattice problem differs from the single-site problem in that coherent states (in solid-state physics terms, Bloch states) can be formed. In other words, the electrons can coherently propagate through the lattice in the same way that wave propagation is possible through coherent diffraction in any periodic set of scatterers. This picture has many attractive features. The coherence of the electronic state causes the resistivity to vanish, thus explaining the low residual resistivity quoted earlier. Secondly the energy scale on

which this coherence occurs would be the single-site Kondo energy, which is determined by the weak hybridization process and is therefore small (typically 10–100 K). Because all the entropy associated with the degeneracies of the f-electron configurations is on the scale of the small Kondo energy, the coefficient γ is much larger than usual and can reach the desired value of $\sim 1 \text{ J mol}^{-1} \text{ K}^{-2}$. The magnetic susceptibility is correspondingly large because the opposition to the development of a magnetic moment in the ground state has as its characteristic energy the small Kondo energy.

There are, however, objections to describing these systems as simple collections of single-site Kondo impurities. The description lacks a clear microscopic derivation, and when one tries such a derivation the interaction effects between the sites do not seem to be negligible. An alternative approach is to consider that the essence of the lattice problem is the formation of the coherent Bloch states and that this comes about through the transfer of f-electrons from site to site. However, such transfer processes are severely inhibited by the strong Coulomb intra-site interaction which forbids extra f-electrons on a site. Therefore the only way to obtain f-electron transfer is to remove some f-electrons from the f-levels and place them in the higher energy states near the Fermi energy. This costs energy but it allows coherent hybridization processes to take place, which gain energy. The balance between these two processes leads to a small number of f-electrons being promoted to the higher energy states and a characteristic energy for the coherent state which has a form similar to the single-site Kondo energy. Again χ and γ are determined by this small coherent energy and are large, but the ratio χ/γ may now be larger than the value given by equation (1). A feature common to all approaches is that the large γ values arise from the entropy associated with the degeneracy of the f-configurations or, in other words, the spin degeneracy of the f-states. Another way of describing this feature is that spin fluctuations of the f-electrons are the dominant electronic excitations responsible for the large values of χ and γ at low temperatures.

There is another possible ground state for the lattice problem, namely a state with integral f-electron atomic configurations and ordered local moments on the f-sites. The energy comparison between such a state and the heavy-electron state is an open theoretical question and until it is solved no theoretical predictions can be made as to which of these intermetallic compounds are the more usual antiferromagnetic metals and which are heavy-electron metals.

Superconductivity. There is one condensed-matter system with strong spin fluctuations and a superfluid state: this is liquid ^3He . (Note that ^3He atoms have spin-1/2 and are fermions, whereas ^4He atoms have spin 0 and are bosons. The ^3He ground state is called superfluid rather than superconducting simply because ^3He atoms are neutral.) Liquid ^3He solidifies at a modest pressure, so that in the liquid state the ^3He atoms are almost localized and therefore can flip their spins easily. In ^3He the superfluid state is known to be an anisotropic p-state (that is, the Cooper pairs of ^3He atoms responsible for the superfluid state are in a relative p-state). It has been shown theoretically that the interactions mediated through spin fluctuations favour p-state superfluidity²³. Also a relative p-state means a much reduced overlap between the fermions and so is favoured when there is a very strong short-range repulsion as is the case for ^3He atoms. The analogy has been made by many authors on both grounds (that is, strong spin fluctuations and strong short-range repulsion) for the heavy-electron systems. In normal metals there is an attractive interaction, mediated by the exchange of phonons, between electrons in states within a Debye energy, $k_B\theta_D$, of the Fermi energy, where θ_D is the Debye temperature. The repulsive Coulomb interaction is reduced for these states because of the mismatch of the energy scale of the Coulomb energy, which is E_F and $k_B\theta_D$. Typically $k_B\theta_D/E_F \approx 10^{-2}$, but in the heavy-electron metals $k_B\theta_D \geq E_F$ so there is no mismatch. Nevertheless

some authors continue to argue for the dominance of phonon-mediated interactions, and therefore for an isotropic or relative s-state pairing. The microscopic theory of the electron pairing necessary to obtain superconductivity in these systems is as yet only qualitative.

Much effort has been put into obtaining a proper description of p-state superconductivity in these crystalline materials which also have important spin-orbit interactions. By the use of group-theoretical techniques the allowed forms can be correctly specified, and there are many possible states which differ in the detailed form of the energy gap which is associated with the electron pairing. In usual s-state pairing the gap is simply isotropic in $\mathbf{k}$ -space and the electrons are in a relative spin-singlet state. However the global antisymmetry of fermions requires non-trivial structure for the gap in other pairing states where electrons are in spin-parallel or triplet states. As a result, in many of these states, the energy gap vanishes along specific directions in $\mathbf{k}$ -space.

Clearly the very-low-energy excitations in superconductors with isotropic gaps will differ from those in superconductors with anisotropic gaps which vanish in certain directions. In the former, properties such as specific heat are exponential as $T \rightarrow 0 \text{ K}$, whereas in the latter they exhibit power-law behaviour (as shown in Fig. 4) as $T \rightarrow 0 \text{ K}$. The observation of such power-law behaviour, as discussed above (Fig. 4), is taken by many as evidence for anisotropic superconductivity. Unfortunately we lack a decisive way to determine the detailed symmetry of the superconducting state: the equivalent of X-ray and neutron scattering for lattice and magnetic transitions does not exist here. Therefore we must proceed by inference, which can lead, and indeed has led, to differing interpretations.

Conclusions

It is clear from the above that the heavy-electron metals represent an exciting new class of materials. They do not fit into the traditional classifications of materials with partially filled atomic shells; namely, metals, semiconductors or Mott insulators. At high temperatures they are essentially indistinguishable from a host of other rare-earth or actinide intermetallic compounds with local moments arising from the partially filled f-shells. But as the temperature is lowered, instead of forming magnetically ordered structures as is usually the case, these materials unexpectedly form an itinerant or delocalized metallic state. The electrons in this state have a much compressed energy scale and correspondingly high effective mass. The theoretical understanding of how such a state arises out of the f-electrons with their strong Coulomb interactions is today a very active problem and will undoubtedly lead to new insights into the general many-body problem. In particular we would like to know what the material parameters are which determine whether, and under what conditions, a material exhibits this heavy-electron state; we would also like to understand the nature of the residual interactions which characterize this metallic state. It is clear from the discussion above that the heavy-electron metallic state can display the phase transitions seen in many other traditional metals; namely, superconductivity (as in UBe_{13}) and weak (or itinerant) magnetism (as in U_2Zn_{17}). But, as we have stressed, there are many intriguing signs that the superconductivity is not simply the usual isotropic (or s-state) kind. As we have also stressed, this anisotropic superconductivity is expected theoretically when the mechanism responsible for the superconductivity is not the usual electron-phonon interaction. Thus it appears that at last such superconductivity has been found in metals. The confirmation of a second mechanism for superconductivity will spur the search for yet other mechanisms. One of the motivations in this long search was the hope that by achieving superconductivity through unconventional mechanisms, new and possibly higher energy scales would apply, leading to higher values of T_c . This has not yet been obtained, but two important areas must continue to be explored. Materials scientists must

continue the search for a higher T_c , while theorists strive for an improved understanding of the factors which determine and limit T_c . Even modest rises in T_c in these materials might have important technical consequences in the generation of very high magnetic fields, as their superconducting state is extremely stable with respect to magnetic field intensity.

H.R.O. and T.M.R. thank the Center for Materials Sciences at Los Alamos for hospitality. Work done at Los Alamos was under the auspices of the U.S. Department of Energy. Partial financial support from the Schweizerische Nationalfonds zur Förderung der wissenschaftlichen Forschung is gratefully acknowledged.

1. Stewart, G. R. *Rev. mod. Phys.* **56**, 755-787 (1984).
2. Andres, K., Graebner, J. E. & Ott, H. R. *Phys. Rev. Lett.* **35**, 1779-1782 (1975).
3. Steglich, F. *et al. Phys. Rev. Lett.* **43**, 1892-1896 (1979).
4. Ott, H. R., Rudigier, H., Fisk, Z. & Smith, J. L. *Phys. Rev. Lett.* **50**, 1595-1598 (1983).
5. Ott, H. R., Rudigier, H., Delsing, P. & Fisk, Z. *Phys. Rev. Lett.* **52**, 1551-1554 (1984).
6. Stewart, G. R., Fisk, Z., Willis, J. O. & Smith, J. L. *Phys. Rev. Lett.* **52**, 679-682 (1984).
7. Maple, M. B. & Fischer, Ø. (eds) *Superconductivity in Ternary Compounds II* (Springer, Berlin, 1982).
8. Bardeen, J., Cooper, L. N. & Schrieffer, J. R. *Phys. Rev.* **108**, 1175-1204 (1957).
9. Leggett, A. J. *Rev. mod. Phys.* **47**, 331-414 (1975).
10. Ott, H. R. *et al. Phys. Rev. Lett.* **52**, 1915-1918 (1984).
11. Bishop, D. J. *et al. Phys. Rev. Lett.* **53**, 1009-1011 (1984).
12. MacLaughlin, D. E. *et al. Phys. Rev. Lett.* **53**, 1833-1836 (1984).
13. Golding, B. *et al. Phys. Rev. Lett.* **55**, 2479-2482 (1985).
14. Maple, M. B. *et al. Phys. Rev. Lett.* **54**, 477-480 (1985).
15. Ott, H. R., Rudigier, H., Fisk, Z. & Smith, J. L. *Phys. Rev. B* **31**, 1651-1652 (1985).
16. Batlogg, B. *et al. Phys. Rev. Lett.* **55**, 1319-1323 (1985).
17. Cox, D. E. *et al. Phys. Rev. B* (in the press).
18. Varma, C. M. *Comments Solid St. Phys.* **11**, 221-240 (1985).
19. Kasuya, T. & T. Saso (eds) *Theory of Heavy Fermions and Valence Fluctuations* Proc. 8th Taniguchi Symp. (Springer, Berlin, 1985).
20. Lee, P. A., Rice, T. M., Serene, J. W., Sham, L. J. & Wilins, J. W. *Comments Con. Matter Phys.* **12**, 99-130 (1986).
21. Anderson, P. W. *Phys. Rev.* **124**, 41-53 (1961).
22. Tsvetick, A. M. & Wiegmann, P. B. *Adv. Phys.* **32**, 453-713 (1983).
23. Anderson, P. W. & Brinkman, W. F. *Phys. Rev. Lett.* **30**, 1108-1111 (1973).

ARTICLES

Molecular stratigraphy: a new tool for climatic assessment

S. C. Brassell*, G. Eglinton*, I. T. Marlowe*, U. Pflaumann† & M. Sarnthein†

* Organic Geochemistry Unit, University of Bristol, School of Chemistry, Cantock's Close, Bristol BS8 1TS, UK

† Geologisch-Paläontologisches Institut und Museum, Universität Kiel, Olshausenstrasse 40/60, D-2300 Kiel, FRG

Variations in sea-surface temperatures over the past 500,000 years are inferred from the relative abundance behaviour of two organic compounds, C_{37} alkenones over the upper 8 metres of a sediment core from the eastern equatorial Atlantic. This molecular record, ascribed to contributions from prymnesiophyte algae, correlates well with the variations in the $\delta^{18}O$ signal for the calcareous skeletons of certain planktonic foraminifera, thus providing the first demonstration of a new stratigraphical technique, which may be especially valuable where methods based on carbonate $\delta^{18}O$ fail.

THE ubiquitous unicellular marine coccolithophorid *Emiliania huxleyi* (Prymnesiophyceae)¹ contains long-chain (C_{37} - C_{39}) di-, tri- and tetra-unsaturated methyl and ethyl ketones (alkenones)²⁻⁴ whose unsaturation changes with growth temperature in laboratory cultures⁵. These compounds also occur in contemporary bottom sediments^{6,7}, presumably derived from phytoplankton in the photic zone of the overlying water column⁸. Differences in the alkenone unsaturation of contemporary sediments from two different climatic regimes⁹ suggest that this feature might reflect, and thereby provide a measure of, water temperatures in the euphotic zone¹⁰. Accordingly, the distribution of alkenones in many Quaternary marine sediments from various latitudes of different mean sea-surface temperatures (SST) has been examined. We now report the alkenone unsaturation in a gravity core from the eastern equatorial Atlantic ocean, covering the past million years, and compare this record with that of the $\delta^{18}O$ of planktonic foraminifera, an established indicator of a combination of global ice sheet variations and local SST in the euphotic zone^{11,12}.

Organic compounds

Improved chromatographic and spectrometric methods for the analysis of complex mixtures have greatly extended the number and variety of organic compounds identified in geological materials¹³ and have aided the better understanding of their origins and sedimentary fate^{13,14}. Often, such marker compounds in bottom sediments can be directly related to their source organisms¹⁴; a record that can survive passage through food web processes⁸ and, subsequently, consolidation and the effects

of diagenesis¹⁵. Such signals best escape disturbance by post-depositional heterotrophic (notably microbial) activity¹⁶ in areas of high sediment accumulation rates and low geothermal gradient¹⁵, as in the deep sea.

The search for organic compounds that are diagnostic of their biological origins in marine sediments¹⁷ and sediment traps¹⁸ has led to the recognition of specific lipid markers for many different types of organisms, including algae such as dinoflagellates^{19,20}, bacteria, notably methanogens²¹, and terrestrial higher plants, such as conifers²². In addition, there is the possibility that their distribution patterns may depend on growth conditions; indeed, aquatic organisms can biosynthetically tailor their lipid composition to match environmental conditions of stress²³. Specifically, they maintain the fluidity of their membranes by changing the molecular composition of the lipid bilayer, either in chain length or in unsaturation²³. Laboratory culturing experiments show changes in lipid unsaturation, notably for carboxylic acids²⁴ and wax esters²⁵, in response to temperature variations. These observations are likely to reflect the behaviour of phytoplankton populations in the natural environment, which may also be preserved in the underlying sedimentary record.

Long-chain alkenones

The less labile constituents of marine phytoplankton include series of long-chain (n - C_{37} to n - C_{39}) alkenones^{2,7} which appear to be restricted to a few species of the class Prymnesiophyceae, notably coccolithophorids^{3,4} of the family Gephyrocapsaceae. The unsaturation of such compounds shows a temperature dependence (more at lower temperatures) in laboratory cultur-

Table 1 Alkenone unsaturation index (U_{37}^K) values for Quaternary marine sediments and for sediment trap material

Site*	Location	Codes	Samples	No.	U_{37}^K	Ref.
Sediments						
A	North Sea	IGS SLN33, SP333CS		2	0.40–0.51	2
B	Porcupine Seabight	D51704		1	0.54	6†
C	Japan Trench	DSDP 440		4	0.51–0.55	2, 47†
D	Ionian Sea	10103		5	0.44–0.79	52
E	Cretan Basin	DSDP 378		1	0.66	53
F	Messina Abyssal Plain	DSDP 374		1	0.47	53
G	San Miguel Gap	DSDP 467		1	0.58	54
H	Orca Basin	DSDP 618		1	0.84	†
I	Gulf of California	DSDP 479, 481		3	0.86–0.95	†
J	Middle America Trench	DSDP 487, 491, 493		5	0.78–0.98	9†
K	Middle America Trench	DSDP 494, 495, 496		3	0.84–0.92	†
L	Cariaco Trench	DSDP 147		1	0.89	†
M	Kane Gap	M16415-2		72	0.76–0.94	This paper
N	Sierra Leone Rise	M13519		3	0.88–0.97	†
O	Peruvian Shelf	110		1	0.87	55
P	Peruvian Shelf	BC7		5	0.23–0.70	44
Q	Walvis Ridge	DSDP 532		1	0.56	17
R	Walvis Bay	IOS 1/75		2	0.52	2
S	Corner Inlet	—		1	0.56	56
T	South Atlantic	DSDP 513		1	0.33	†
Sediment traps						
P	Peruvian Shelf	—		8	0.26–0.56	43

Codes refer to sample sites (DSDP 440 denotes Deep Sea Drilling Project Site 440) as specified in the refs. U_{37}^K denotes the alkenone unsaturation index. The terminology [U = unsaturation, K = ketone (alkenone), 37 = chain length of ketone] will allow for future use with different molecular parameters.

$$U_{37}^K = \frac{[C_{37:2}] - [C_{37:4}]}{[C_{37:2} + C_{37:3} + C_{37:4}]}$$

where, for example, $C_{37:2}$ represents the quantity of the C_{37} alkadienone, as measured by GC peak integration. U_{37}^K can have values between -1.0 (all $C_{37:4}$) and $+1.0$ (all $C_{37:2}$), but varies between $+0.23$ and $+0.98$ in Quaternary marine sediments, and between -0.2 and $+1.0$ for marine sediments of Albian to Holocene age^{6,28}.

* Letters A to T relate to sample latitudes shown in Fig. 1.

† I.T.M., S.C.B. and G.E., unpublished data.

ing experiments with *E. huxleyi* and *Isochrysis galbana*⁵. The distribution patterns of these compounds appear largely unaltered by food web processes, surviving digestion by copepods⁸ and mussels²⁶. Alkenones are easily recovered from sediments by routine extraction procedures and can be readily recognized at subnanogram levels by capillary gas chromatography (GC) and gas chromatography/mass spectrometry (GC/MS)^{2,6}.

Since their first description²⁷, these alkenones have been found in numerous Quaternary marine sediments^{6,7} (Table 1), although they vary in their unsaturation. Here, this variation is expressed as an alkenone unsaturation index, U_{37}^K (Table 1), defined as:

$$U_{37}^K = [C_{37:2}] - [C_{37:4}] / [C_{37:2} + C_{37:3} + C_{37:4}]$$

For most sediments the absence of $C_{37:4}$ simplifies U_{37}^K to $[C_{37:2}] / [C_{37:2} + C_{37:3}]$ which varies inversely with unsaturation. Replicate GC analyses suggest that U_{37}^K values are accurate to ± 0.02 . In Quaternary marine sediments (Table 1), the greater values for U_{37}^K , which we interpret as corresponding to higher water temperatures, are seen in low latitudes, and the lowest values for U_{37}^K are those for the highest latitudes (Fig. 1). This broad relationship suggests mutual trends due to the inherent global and regional variations in SST resulting from oceanic circulation features, such as gyres and upwelling of cooler intermediate water masses in subtropical latitudes. The overall correspondence, however, suggests that alkenone unsaturation in oceanic bottom sediments is an indirect measure of SST. The same series of alkenones has also been recognized in sediments of Pliocene, Miocene, Oligocene and Eocene age^{6,7} and, recently, in Cretaceous black shales²⁸.

Kane Gap sediments

The direct assessment of water temperatures (and hence palaeoclimates) from molecular data requires testing by relating U_{37}^K to an established technique, namely the oxygen isotope ($\delta^{18}O$) signal for the carbonate tests of foraminifera. Herein, the correlation between the alkenone distributions (as U_{37}^K values) and $\delta^{18}O$ ($CaCO_3$) values for both planktonic and benthic foraminifera, is explored for a 13-m sediment core from near the Kane Gap, eastern equatorial Atlantic (Fig. 2). The core (M16415-2), collected during *Meteor* cruise 65 in 1983 (ref. 29), possessed many features appropriate for this study. First, it provided an almost complete record of glacial/interglacial cycles over the past million years. Second, examination of a gravity core (M13519) from the nearby Sierra Leone Rise³⁰ (Table 1; Fig. 2) revealed sufficient quantities of alkenones to permit their evaluation by GC (Fig. 3) in 2–3-cm sections of sediment cores from this region. Third, the oxygen isotope records of the planktonic and benthic foraminifera were determined, using a new MAT 251 instrument capable, for the first time, of measuring single specimens (one or two), using modified standard procedures^{12,31} from throughout the sequence, except where precluded by carbonate dissolution. When combined with magnetic stratigraphy, oxygen isotope stratigraphy provides a high-resolution timescale for the past 3 Myr, correlated on a global scale^{32,33}.

The down-core variations in oxygen isotope compositions of the three species of foraminifera and in alkenone unsaturation (Fig. 4) show generally similar trends for the glacial/interglacial

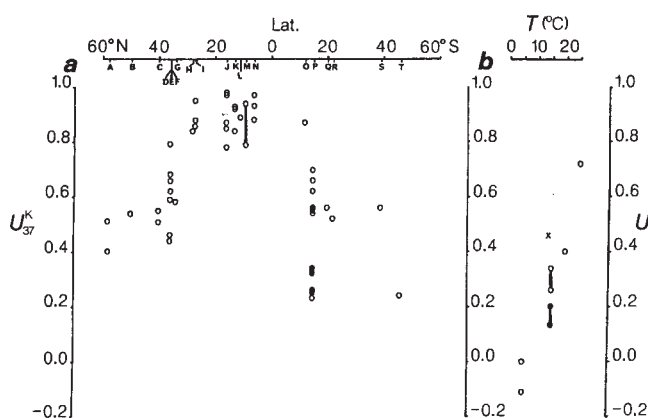


Fig. 1 Alkenone unsaturation index (U_{37}^K) values (Table 1) for: **a**, Quaternary marine sediments at sites A to T (○) and sediment traps (●) plotted against latitude; and **b**, laboratory cultures of *E. huxleyi*⁵ strain 92 (○) and strain 92d (●) and a plankton tow (×) off Cape Finisterre, plotted against water temperature. The vertical bars indicate where numerous samples fall within that particular range of values of U_{37}^K . The sediment trap data in **a**, which display a considerable range of values, would become a single, time-averaged value on sedimentation.

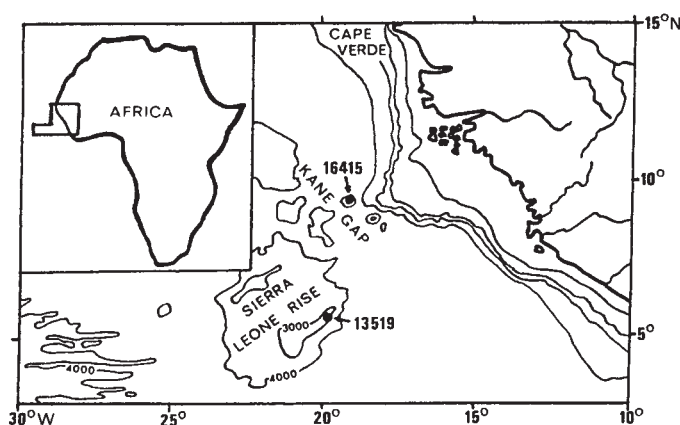


Fig. 2 Locations of Meteor site M13519 on the Sierra Leone Rise and site M16415 in the Kane Gap (partly redrawn from ref. 28). Water depths are in metres.

stages, that is, higher values of $\delta^{18}\text{O}$, corresponding to cooler glacial episodes, are observed at those horizons with lower values of U_{37}^K , and vice versa. The coccolithophorids held to be the principal organisms biosynthesizing the alkenones in the oceans inhabit the upper 150 m of the water column³⁴, although they can occur at greater depths³⁵. Thus, the $\delta^{18}\text{O}$ record for the planktonic foraminifera *Globigerinoides sacculifer* and *Globigerinoides ruber* should parallel that of alkenone unsaturation. These species occur in tropical warm waters where the $\delta^{18}\text{O}$ values for their carbonate tests reflect ice volumes and local SST³⁰, with the former more significant. By contrast, the benthic foraminifer, *Cibicides wuellerstorfi*, inhabits deeper water masses and its $\delta^{18}\text{O}$ principally reflects the waxing and waning of ice sheets rather than sea temperatures³⁰; it is therefore less relevant to U_{37}^K than the planktonic foraminifera. Furthermore, its sampling intervals are greater and the effects of its ecological disappearance are more pronounced, being significant as recent as Stage 6 (Fig. 4). Similarly, $\delta^{18}\text{O}$ data for *G. ruber* are less complete than those for *G. sacculifer*, chosen for detailed comparison with the alkenone unsaturation.

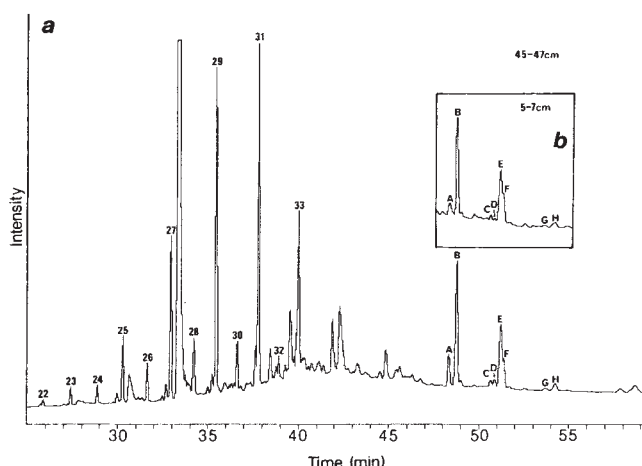


Fig. 3 **a**, Gas chromatogram of total extract (CH_3OH , then $\text{CH}_2\text{Cl}_2/\text{CH}_3\text{OH}$ with sonication) of sample collected at 45–47 cm depth interval from core M16415-2 recovered from the Kane Gap (Fig. 2). Numbered peaks are *n*-alkanes: the number indicates the chain length; those designated A–H are alkenones^{2–7}. A, heptatriaconta-8,15,22-trien-2-one [$\text{CH}_3(\text{CH}_2)_{13}\text{CH}=\text{CH}(\text{CH}_2)_5\text{CH}=\text{CH}(\text{CH}_2)_5\text{CH}=\text{CH}(\text{CH}_2)_5\text{COCH}_3$]; B, heptatriaconta-15,22-dien-2-one [$\text{CH}_3(\text{CH}_2)_{20}\text{CH}=\text{CH}(\text{CH}_2)_5\text{CH}=\text{CH}(\text{CH}_2)_5\text{COCH}_3$]; C, octatriaconta-9,16,23-trien-3-one; D, octatriaconta-9,16,23-trien-2-one; E, octatriaconta-16,23-dien-3-one; F, octatriaconta-16,23-dien-2-one; G, nonatriaconta-10,17,24-trien-3-one; H, nonatriaconta-17,24-dien-3-one. The relative proportions of $\text{C}_{37:3}$ and $\text{C}_{37:2}$ alkenones (peaks A and B, respectively) are typical of the lower U_{37}^K values (0.78) for core M16415-2. **b**, Partial gas chromatogram of total extract of 5–7 cm depth interval from core M16415-2, illustrative of the higher U_{37}^K values (0.89) determined in this study. GC conditions: Carlo Erba 4160 instrument with on-column injection, fitted with 50-m CPSil 5CB flexible silica WCOT capillary column, temperature programmed from 60 to 300 °C at 6 °C min^{-1} . U_{37}^K values were calculated from the peak areas of A and B, determined using a VG Minichrom data system. Compound identifications were confirmed by GC/MS (Carlo Erba Mega chromatograph equipped with 25-m OV-1 flexible silica column, programmed as above, and coupled to a Finnigan 4000 mass spectrometer). Data acquisition by INCOS 2300 data system, scanning m/z 50–600 at 1-s intervals.

$\delta^{18}\text{O}$ for *G. sacculifer* and U_{37}^K correlate well ($r=0.932$; Fig. 5a) to almost 2 m and are broadly parallel ($r=0.676$; Fig. 5b) to 8 m. Below 8 m they show little agreement. The discrepancies in Fig. 4 may stem from the low sampling frequency due to CaCO_3 dissolution, from bioturbation and core disturbance, or from the different intervals of $\delta^{18}\text{O}$ and alkenone determinations. The excellent correlation between $\delta^{18}\text{O}$ for *G. sacculifer* and U_{37}^K for the upper 2 m is striking, and includes the temperature profile characteristic of a gradual cooling and subsequent rapid warming over the interglacial/glacial cycle^{11,30,36} of the past 120,000 yr. A general correspondence between the two data sets holds to 8 m, that is the past 550,000 yr. Such similarity of $\delta^{18}\text{O}$ and U_{37}^K data suggests a common causality, namely the systematic variations in the Earth's orbit, described by Milankovitch's theory^{36,37}. Such orbital changes affect global ice volumes and the temperature of oceanic water masses, which are both reflected in the oxygen isotope composition of planktonic foraminifera, such as *G. sacculifer*, whereas the alkenones produced by phytoplankton should reflect SST, but not any direct effects due to changes in global ice volumes. Fluctuations in alkenone unsaturation should therefore, be an indirect manifestation of the Earth's orbital cycles^{33,38}. The discrepancies in the U_{37}^K and $\delta^{18}\text{O}$ profiles raise the intriguing prospect that interpretation of their combination may constitute a general method for assessing the relative contributions of ice

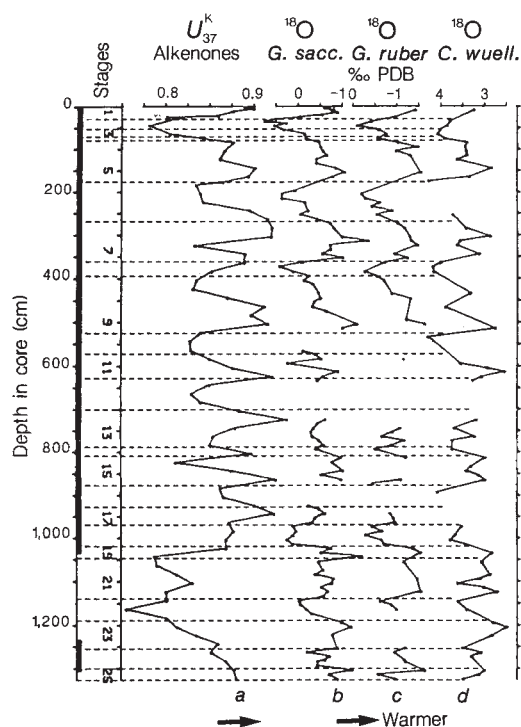


Fig. 4 Downhole plots of the direct measures of alkenone unsaturation index (U_{37}^K) for sediment extracts and of $\delta^{18}\text{O}$ values for foraminifera in Meteor core 16415-2. *a*, Alkenone unsaturation index (U_{37}^K ; Table 1). *b-d*, $\delta^{18}\text{O}$ values of CaCO_3 of foraminifera: planktonic *G. sacculifer* and *G. ruber* and benthic *C. wuellerstorfi*. Gaps between the $\delta^{18}\text{O}$ data indicate complete dissolution of CaCO_3 . The sampling intervals for *d* are greater than those for *a*, *b* and *c*. The oxygen isotope stages³⁰ are indicated by numbers between the ordinates and their boundaries are designated by broken lines (Stage 3 is probably reduced due to an hiatus). The magnetic stratigraphy (thickening of ordinate) for the core is that determined by U. Bleil (personal communication). The general trends reflecting cooler and warmer temperatures are indicated.

volume and SST within the oxygen isotope signal. Indeed, the direct correspondence between U_{37}^K and $\delta^{18}\text{O}$ values to 2 m suggests that $\delta^{18}\text{O}$ values may reflect SST more closely than originally thought—at least in this part of the ocean.

The behaviour pattern of U_{37}^K below 8 m (down to 1 Myr BP) differs significantly from that above 8 m. Diagenetic alteration of alkenone distributions and/or changes in the processes affecting alkenone degradation before burial, are unlikely to be responsible; a transition in the source organisms of the alkenones seems more probable. Such factors cannot be distinguished at present, but a possible source species³⁹, the coccolithophorid *Pseudoemiliania lacunosa*, became globally extinct at a time^{40,41} corresponding to a depth of ~6.5 m within this core, and *E. huxleyi* first appeared <270,000 yr BP (at ~4 m)⁴¹. Some ancestor/predecessor species of *E. huxleyi*, such as *Gephyrocapsa margaritella*⁴⁰, must show similar alkenone compositional behaviour back to at least 550,000 yr BP (that is, 8 m). If variations in alkenone unsaturation are indeed related to membrane fluidity and hence temperature, then this behaviour should be a general feature of such organisms. The extent of alkenone unsaturation of different species, however, need not be uniform at a given temperature.

U_{37}^K temperature calibration

The molecular record of SST requires a defined link between the alkenone unsaturation index for relevant organisms and that for bottom sediments. There is a need to evaluate U_{37}^K for natural

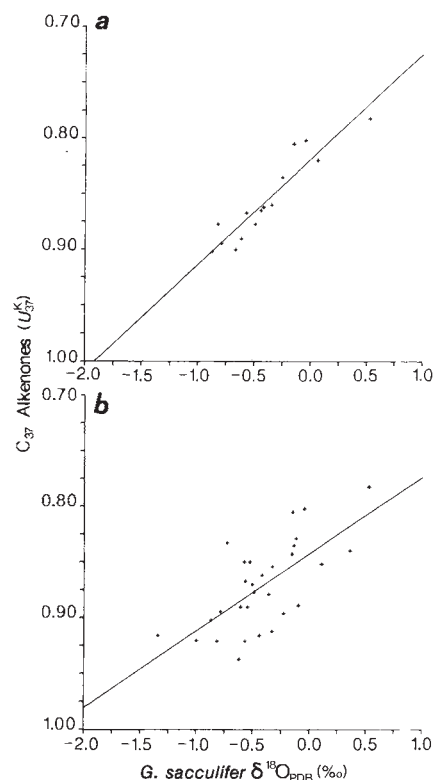


Fig. 5 Regression plots of the alkenone unsaturation index (U_{37}^K) for sediment extracts against $\delta^{18}\text{O}$ for the carbonate of *G. sacculifer*. *a*, Interpolated data for 5–185 cm depth in Meteor core M16415-2; $r = 0.932$. *b*, Original data, 0–800 cm depth in Meteor core M16415-2; $r = 0.676$.

populations of phytoplankton, principally coccolithophorids such as *E. huxleyi*, collected from present-day oceanic water masses of known temperature at depths comparable to those where appropriate planktonic foraminifera produce their tests, and at times of maximum primary productivity. In the eastern equatorial Atlantic, the standing crop of *E. huxleyi* is probably highest during the winter months, by analogy with waters around Bermuda⁴². A plankton tow taken off Cape Finisterre (43°11' N, 09°33' W) in early July, from waters of 14 °C containing *E. huxleyi* and *Coccolithus pelagicus*, yielded small amounts of alkenones with a U_{37}^K of 0.46 (Fig. 1); a value comparable to those of sediments from similar latitudes. A broader base of alkenone data, however, exists for laboratory cultures of *E. huxleyi* grown at several temperatures⁵ (Fig. 1), although these show discrepancies compared with those for sediments. For example, tetra-unsaturated alkenones (alkatetraenones) occur in the cultured algae^{4,6} at all temperatures⁵, but are rarely detected in oceanic sediments⁶, even when U_{37}^K is <0.4 (Table 1). Indeed, U_{37}^K values of *E. huxleyi* grown at a given temperature are lower (Fig. 1) than those of sediments deposited beneath waters of similar temperature. However, the strains of *E. huxleyi* in culture were collected originally from the south-west approaches to the English Channel (~50 °N; J. C. Green, personal communication) and their tolerance to higher growth temperatures (20 or 25 °C) may be limited. The influence of sedimentation processes on U_{37}^K values, where zooplankton faecal pellets may facilitate rapid transportation of alkenones to bottom sediments⁸, has not been tested. The amounts of alkenones in the plankton tow off Cape Finisterre suggest that they may have been derived, in part, from semi-digested phytoplankton in the guts of zooplankton within this sample (I.T.M., S.C.B. and G.E., unpublished data). Certainly, the range of U_{37}^K values in sediment traps from Peru⁴³ (Table 1; Fig. 1) requires

explanation, and the low U_{37}^K value for the surface sediment in this locality⁴⁴ is also unexpected. The sedimentary record, however, represents a time-averaged signal, so that short-term fluctuations and extremes in U_{37}^K values may make little ultimate impact.

The U_{37}^K values of sediments from the Kane Gap (Fig. 4) where SST remain at $\sim 26^\circ\text{C}$ throughout the year^{45,46} are greater than those of *E. huxleyi*⁵ grown at temperatures up to 25°C in the laboratory. For M16415-2 the extreme values of U_{37}^K differ by 0.18 over the entire core sequence and by 0.12 over the last warming episode, that is, Termination I. Laboratory cultures of *E. huxleyi* (Fig. 1) indicate that the latter value corresponds to a temperature change of $2\text{--}5^\circ\text{C}$. Despite the obvious uncertainties in such calibration based on culture experiments this value is comparable to the 4°C reported for Termination I, based on February SST values for 18,000 yr BP and for the present day, estimated from foraminiferal assemblages⁴⁵.

Additional influences on U_{37}^K values will include the geological and evolutionary succession of species^{40,41} that contributed alkenones to sediments³⁹. Clearly, species-related effects on the distributions of alkenones need to be assessed, yet their biosynthesis is likely to be more conservative than the morphology of the algae. Diagenetic effects may also affect the U_{37}^K record, but the notable absence of mono-unsaturated homologues in all sediments that contain alkenones suggests that chemical reduction of alkenones does not occur, and argues against any effects on U_{37}^K stemming from diagenetic conversion of alkatrienones into alkadienones.

Molecular geochemical indicators

Future applications of molecular geochemical indicators of water temperatures will include evaluation of sediment sequences deposited below the calcite compensation depth (CCD), where planktonic foraminiferal tests are not preserved. In core

M16415-2, alkenone data cover intervals lacking such tests, due to dissolution, for example, in Stage 12 (Fig. 4). In addition, alkenones occur as abundant constituents of diatomaceous oozes from the Japan Trench (Table 1) deposited below the CCD^{2,47,48} and, consequently, bereft of calcareous foraminiferal tests. The alkenone unsaturation index may be a uniquely useful stratigraphical tool in such instances. Alkenones have also been detected in lacustrine sediments⁴⁹, with U_{37}^K values of $+0.16$ to -0.58 , and could therefore serve as palaeoclimatic indicators in such environments. Future improvements in alkenone detection limits will reduce sample sizes and thereby increase stratigraphical resolution—although this, inevitably, is limited by the extent of bioturbation.

The alkenones are independent of isotopic factors, although subject to effects related to biochemical evolution and diagenesis. They are a single type of organic compound among the thousands known to occur in sediments^{13,14}. We anticipate that U_{37}^K is the forerunner of a substantial number of molecular parameters⁵⁰ for assessment of environmental conditions of sediment deposition⁵¹. To this end, we are now exploring correlations between molecular and other data using principal-component and spectral analysis and are investigating evidence for Milankovitch cycles in sedimentary alkenone data⁵⁰.

We thank the NERC (GR3/2951 and GR3/3758) for support and a studentship (I.T.M.), and the Royal Society for funds towards the VG Minichrom system. We thank Dr J. C. Green (MBA, Plymouth) for collaboration with studies of alkenones in living algae and Dr H. Erlenkeuser (Kiel University ^{14}C laboratory) for access to new preparation facilities of a MAT 251 mass spectrometer for measurements of foraminiferal $\delta^{18}\text{O}$. Professor U. Bleil (Bochum) provided unpublished magnetostratigraphy for core M16415-2. We also thank Drs R. G. Brereton and J. Grimalt for helpful discussions during the final preparation of this paper and other colleagues at Bristol for access to unpublished data.

Received 14 August; accepted 30 December 1985.

- Hay, W. W. & Mohler, H. P. *Trans. Gulf-Coast Ass. geol. Soc.* (eds Hay, W. W. et al.) **17**, 426–480 (1967).
- Volkman, J. K., Eglinton, G., Corner, E. D. S. & Sargent, J. R. in *Advances in Organic Geochemistry 1979* (eds Douglas, A. G. & Maxwell, J. R.) 219–227 (Pergamon, Oxford, 1980).
- Volkman, J. K., Eglinton, G., Corner, E. D. S. & Forsberg, T. E. V. *Phytochemistry* **19**, 2619–2622 (1980).
- Marlowe, I. T. et al. *Br. phycol. J.* **19**, 203–216 (1984).
- Marlowe, I. T., Green, J. C., Brassell, S. C., Eglinton, G. & Course, P. A. *Phytochemistry* (submitted).
- Marlowe, I. T., Brassell, S. C., Eglinton, G. & Green, J. C. *Org. Geochem.* **6**, 135–141 (1984).
- de Leeuw, J. W., van der Meer, F. W. & Rijpstra, W. I. C. in *Advances in Organic Geochemistry 1979* (eds Douglas, A. G. & Maxwell, J. R.) 211–217 (Pergamon, Oxford, 1980).
- Volkman, J. K., Corner, E. D. S. & Eglinton, G. in *Colloq. int. CNRS no. 293* (ed. Daumas, R.) 185–197 (Editions CNRS, Paris, 1980).
- Brassell, S. C., Eglinton, G. & Maxwell, J. R. *Init. Rep. DSDP* **66**, 557–580 (1981).
- Brassell, S. C. & Eglinton, G. in *Heterotrophic Activity in the Sea* (eds Hobbie, J. E. & Williams, P. J. leB.) 481–503 (Plenum, New York, 1984).
- Shackleton, N. J. & Opdyke, N. D. *Quat. Res.* **3**, 39–55 (1973).
- Duplessy, J. C. in *Climatic Change* (ed. Gribbin, J.) 46–67 (Cambridge University Press, 1978).
- Mackenzie, A. S., Brassell, S. C., Eglinton, G. & Maxwell, J. R. *Science* **217**, 491–504 (1982).
- Brassell, S. C., Eglinton, G. & Maxwell, J. R. *Biochem. Soc. Trans.* **11**, 575–586 (1983).
- Brassell, S. C. & Eglinton, G. in *Advances in Organic Geochemistry 1981* (eds Bjorøy, M. et al.) 684–697 (Wiley, New York, 1983).
- Boon, J. J., de Leeuw, J. W. & Schenck, P. A. *Geochim. cosmochim. Acta* **39**, 1559–1565 (1975).
- Brassell, S. C. & Eglinton, G. in *Coastal Upwelling: Its Sediment Record* (eds Suess, E. & Thiede, J.) 545–571 (Plenum, New York, 1983).
- Gagosian, R. B., Nigrelli, G. E. & Volkman, J. K. in *Coastal Upwelling: Its Sediment Record* (eds Suess, E. & Thiede, J.) 241–272 (Plenum, New York, 1983).
- Boon, J. J. et al. *Nature* **277**, 125–127 (1979).
- Robinson, N., Eglinton, G., Brassell, S. C. & Cranwell, P. A. *Nature* **308**, 439–441 (1984).
- Brassell, S. C., Wardroper, A. M. K., Thomson, I. D., Maxwell, J. R. & Eglinton, G. *Nature* **290**, 693–696 (1981).
- Simoneit, B. R. T. *Geochim. cosmochim. Acta* **41**, 463–476 (1977).
- Harwood, J. R. & Russell, N. J. *Lipids in Plants and Microbes* (George Allen & Unwin, London, 1984).
- Holton, R. W., Blecker, H. H. & Onore, M. *Phytochemistry* **3**, 595–602 (1964).
- Russell, N. J. & Volkman, J. K. *J. gen. Microbiol.* **118**, 131–141 (1980).
- Rowland, S. J. & Volkman, J. K. *Mar. env. Res.* **7**, 117–130 (1982).

- Boon, J. J., van der Meer, F. W., Schuyt, P. J. W., de Leeuw, J. W. & Schenck, P. A. *Init. Rep. DSDP* **40**, 627–637 (1978).
- Farrimond, P., Eglinton, G. & Brassell, S. C. in *Advances in Organic Geochemistry 1985* (eds Leythaeuser, D. & Rullkötter, J.) (Pergamon, Oxford, in the press).
- Sarnthein, M., Kögler, F. C. & Werner, F. *Geol. Paläontolog. Inst. Univ. Kiel Ber. No. 2* (1983).
- Sarnthein, M., Erlenkeuser, H., von Grafenstein, R. & Schröder, C. *'Meteor' Forsch.-Ergebn.* **C38**, 9–24 (1984).
- Ganssen, G. *'Meteor' Forsch.-Ergebn.* **C37**, 1–46 (1983).
- Shackleton, N. J. & Opdyke, N. D. *Mem. geol. Soc. Am.* **145**, 449–464 (1976).
- Specmap 1984 in *Milankovitch and Climate* (eds Berger, A. & Imbrie, J.) (Reidel, New York, 1984).
- McIntyre, A., Bé, A. W. H. & Roche, M. B. *Trans. N.Y. Acad. Sci. Ser. II* **32**, 720–731 (1970).
- Raymont, J. E. G. *Plankton and Productivity in the Oceans* (Pergamon, Oxford, 1980).
- Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 1121–1132 (1976).
- Milankovitch, M. K. *Serb. Akad. Beogr. Spec. Publ.* **132** (1941). (Transl. by Israel Prog. for Scientific Translations, Jerusalem, 1969).
- Herterich, K. & Sarnthein, M. in *Milankovitch and Climate* (eds Berger, A. L. & Imbrie, J.) 447–466 (Reidel, New York, 1984).
- Marlowe, I. T., Brassell, S. C., Eglinton, G. & Green, J. C. *Oceanologica Acta* (submitted).
- Santleben, C. *Paläont. Z.* **54**, 91–127 (1980).
- Thierstein, H. R., Geitzenauer, K. R., Molino, B. & Shackleton, N. J. *Geology* **5**, 400–404 (1977).
- McIntyre, A. & Bé, A. W. H. *Deep-Sea Res.* **14**, 561–597 (1967).
- Wakeham, S. G., Farrington, J. W. & Gagosian, R. B. *Org. Geochem.* **6**, 203–215 (1984).
- Volkman, J. K., Farrington, J. W., Gagosian, R. B. & Wakeham, S. G. in *Advances in Organic Geochemistry 1981* (eds Bjorøy, M. et al.) 228–240 (Wiley, New York, 1983).
- Gardner, J. V. & Hays, J. D. *Mem. geol. Soc. Am.* **145**, 211–246 (1976).
- Sarnthein, M. et al. in *Geology of the Northwest African Continental Margin* (eds von Rad, U., Hinz, K., Sarnthein, M. & Seibold, E.) 545–604 (Springer, Berlin, 1982).
- Brassell, S. C. et al. *Init. Rep. DSDP* **56/57**, 1367–1390 (1980).
- Brassell, S. C. et al. in *Advances in Organic Geochemistry 1979* (eds Douglas, A. G. & Maxwell, J. R.) 375–392 (Pergamon, Oxford, 1980).
- Cranwell, P. A. *Geochim. cosmochim. Acta* **49**, 1545–1551 (1985).
- Brassell, S. C. et al. in *Advances in Organic Geochemistry 1985* (eds Leythaeuser, D. & Rullkötter, J.) (Pergamon, Oxford, in the press).
- Didyk, B. M., Simoneit, B. R. T., Brassell, S. C. & Eglinton, G. *Nature* **272**, 216–222 (1978).
- Smith, D. J., Eglinton, G. & Morris, R. J. *Phil. Trans. R. Soc.* (in the press).
- Comet, P. A. thesis, Univ. Bristol (1982).
- McEvoy, J. thesis, Univ. Bristol (1983).
- Smith, D. J., Eglinton, G. & Morris, R. J. *Geochim. cosmochim. Acta* **47**, 2225–2232 (1983).
- Volkman, J. K., Gillan, F. T., Johns, R. B. & Eglinton, G. *Geochim. cosmochim. Acta* **45**, 1817–1828 (1981).

Human oestrogen receptor cDNA: sequence, expression and homology to *v-erb-A*

Stephen Green*, Philippe Walter*, Vijay Kumar*, Andrée Krust*, Jean-Marc Bornert*, Patrick Argos† & Pierre Chambon*‡

* Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Faculté de Médecine, 11 rue Humann, 67085 Strasbourg Cédex, France

† European Molecular Biology Laboratory, Postfach 10.2209, Meyerhofstrasse 1, 6900 Heidelberg, FRG

We have cloned and sequenced the complete complementary DNA of the oestrogen receptor (ER) present in the breast cancer cell line MCF-7. The expression of the ER cDNA in HeLa cells produces a protein that has the same relative molecular mass and binds oestradiol with the same affinity as the MCF-7 ER. There is extensive homology between the ER and the erb-A protein of the oncogenic avian erythroblastosis virus.

THE oestrogen receptor (ER) is found in a wide variety of species and is involved in the programming and regulation of gene expression in vertebrate female sex-accessory tissues (for reviews see refs 1 and 2). It has been suggested that the ER is localized predominantly in the nucleus^{3,4} and that, following the binding of oestrogen, there is an increase in its affinity for some promoter elements of oestrogen-regulated genes (refs 2, 5, 6 and references therein). Very little is known about the underlying molecular mechanism.

Significant levels of ER have been detected in more than 50% of human breast cancers. Approximately 70% of these ER-positive tumours respond to anti-oestrogen therapy compared with only ~5% of ER-negative tumours, suggesting a strong correlation between the growth of breast tumours *in vivo* and the presence of the ER⁷. Oestrogens have been shown to be mitogenic in the ER-positive breast cancer cell line MCF-7 (ref. 8). The mechanism by which the binding of oestradiol to the ER results in stimulation of cell division is unknown, but several gene products are known to be induced (ref. 9 and references therein). Their role in promoting cell growth is unknown.

We have described previously the isolation of complementary DNA clones corresponding to the human ER messenger RNA present in MCF-7¹⁰ and we have now isolated and sequenced additional cDNA clones that cover the entire ER mRNA. We report here the sequence of the complete ER cDNA and its expression in HeLa cells, and show that there is extensive homology between the ER cDNA and the *erb-A* gene of the oncogenic avian erythroblastosis virus (AEV).

Sequence of the ER mRNA and gene

We have described previously the isolation of a 2.1-kilobase (kb) ER cDNA clone (λ OR8, see Fig. 1) using synthetic oligo-nucleotides corresponding to ER peptide sequences¹⁰. Sequencing of λ OR8 showed that it contains a large open reading frame of 1,785 nucleotides, sufficient to encode the ER (relative molecular mass (M_r) ~65,000; ref. 10). Since the most 5' ATG is commonly used to initiate translation in eukaryotes¹¹, we expected λ OR8 to correspond to the 5' end of the ER mRNA, which is ~6.2 kb long¹⁰. An oligo(dT)-primed λ gt10 cDNA library, prepared from MCF-7 cell poly(A)⁺ RNA enriched in ER mRNA¹⁰, was screened with λ OR8 and the isolated clones were mapped. A probe corresponding to the 3' end of one of these clones, λ OR15 (see Fig. 1), was used to re-screen the library. This process was continued in order to isolate clones representing the entire ER mRNA (Fig. 1). No clones extending

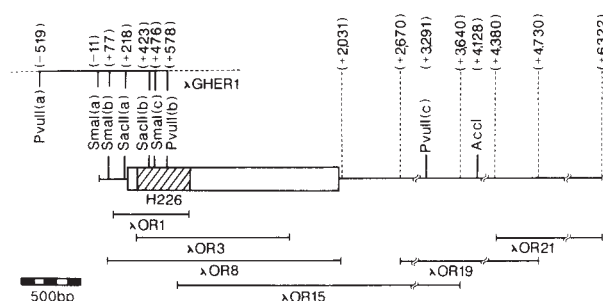


Fig. 1 cDNA clones representing the complete ER mRNA. The MCF-7 ER mRNA is represented by a solid line beginning at nucleotide +1 (the cap site) and ending at nucleotide +6,322. The open reading frame is denoted as a box beginning at nucleotide +233 and ending at nucleotide +2,018. The region which contains the epitope recognized by the MCF-7 ER monoclonal antibody H226 is shown as a hatched box. This region was determined from the overlapping portions of the two λ gt11 cDNA clones (λ OR1 and λ OR3) which react with the H226 antibody¹⁰. Only those restriction enzyme sites referred to in the text are shown. λ GHER1 was isolated from a human genomic library and is shown above the ER mRNA. The region between the cap site and the *PvuII*(b) site at +578 is homologous to the ER mRNA. The dotted lines upstream and downstream of the *PvuII*(a)/*PvuII*(b) region of λ GHER1 represent sequences in the λ clone which have not been analysed. The λ OR8, λ OR15, λ OR19 and λ OR21 clones were isolated from MCF-7 cDNA libraries¹⁰ (also see below).

Methods. An oligo(dT)-primed cDNA library, prepared using sucrose gradient fractions of MCF-7 cell poly(A)⁺ RNA enriched in ER mRNA, was constructed in λ gt10 as described previously¹⁰. The extremities of each of the inserts contain an *EcoRI* site due to the addition of *EcoRI* linkers during the cloning procedure. The insert of λ OR8 was purified on sucrose gradients, nick-translated and used to screen this library. The *PvuII*(c)/*EcoRI* (3' end) fragment of one of the isolated clones, λ OR15, was purified on a 5% polyacrylamide gel, electroeluted, nick-translated and used to rescreen the library which resulted in the isolation of λ OR19. This process was repeated using the 3' *AccI*/*EcoRI* fragment of λ OR19 to obtain λ OR21. The 400-bp *SmaI*(b)/*SmaI*(c) fragment, corresponding to the 5' end of λ OR8, was similarly used to screen a genomic library constructed from human leukocyte DNA partially digested with *Sau3A* and inserted into the *Bam*HI site of λ EMBL3³⁹. The 13-kb insert of one of the clones isolated, λ GHER1, contains a 1.1-kb *PvuII*(a)/*PvuII*(b) fragment (solid line) representing the 5'-flanking sequence and part of the first exon of the ER.

‡ To whom all correspondence should be addressed.

© 1986 Nature Publishing Group

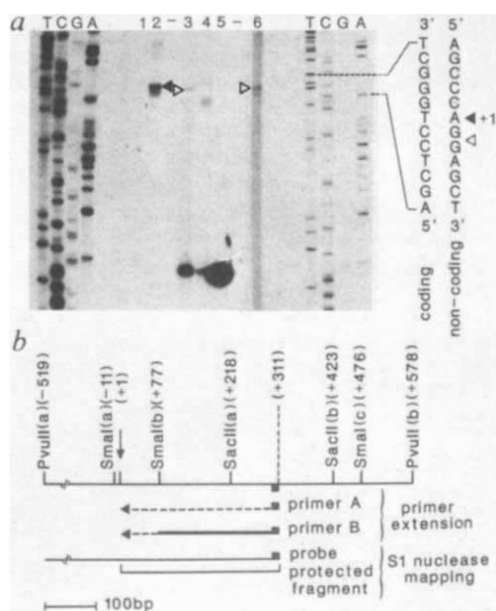


Fig. 3 Mapping of the ER mRNA cap site. *a*, The 1.1-kb *PvuII*/*PvuII* fragment of λ GHER1 (Fig. 1) was subcloned into M13mp18. A single-stranded DNA probe (*b*) corresponding to this genomic fragment was prepared and used for S_1 -nuclease mapping with (lane 2) or without (lane 1) sucrose gradient fractions of MCF-7 poly(A)⁺ RNA enriched in ER mRNA¹⁰. Primer extension mapping was performed using primers A and B (*b*), reverse transcriptase and poly(A)⁺ RNA enriched in ER mRNA. Primer A is a 17-mer oligonucleotide (5'-TCAGGGGCTCCAGCTCG-3') corresponding to positions +295 to +311 of the ER mRNA sequence (Fig. 2); it was elongated and electrophoresed without treatment with S_1 nuclease (lane 6). Primer B (lane 5) is 232 nucleotides long, from the *SmaI*(b) site to position +311. It was extended in the presence (lane 3) or absence (lane 4) of ER mRNA and treated with S_1 nuclease before electrophoresis. The arrowheads show the position of the 5' end of the ER mRNA as defined by either S_1 -nuclease mapping (closed arrowheads) or primer extension (open arrowheads). The four lanes (TCGA) show the sequence of the mRNA coding strand of the *PvuII*/*PvuII* fragment of λ GHER1 (dideoxy sequencing using the 5'-end-labelled primer A).

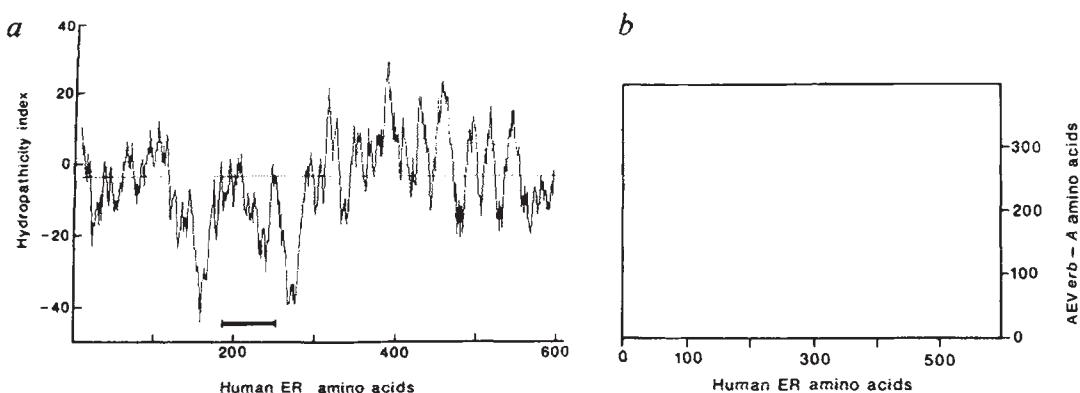
Methods. For both the S_1 -nuclease and primer extension mapping, the 17-mer oligonucleotide, primer A, was 5'-end-labelled with [γ -³²P]ATP and polynucleotide kinase. The S_1 -nuclease mapping was carried out as follows. The labelled primer A was used for second-strand synthesis using the *PvuII*/*PvuII* M13mp18 clone as a template with cold nucleotides (500 μ M) and all other conditions as for dideoxy sequencing. The DNA was cut at the *EcoRI* site in the M13mp18 polylinker and the 830-nucleotide single-stranded probe was purified by polyacrylamide gel electrophoresis. This probe (600,000 c.p.m., 0.12 pmol) was hybridized to sucrose gradient fractions of MCF-7 poly(A)⁺ RNA containing ~500 pg ER mRNA (40 mM PIPES pH 6.5, 400 mM NaCl, 50% formamide, 2 min at 80 °C, 2 min at 100 °C and overnight at 42 °C), and then digested with 60 U of S_1 nuclease (BRL) for 2 h at 25 °C in 50 mM Na-acetate pH 4.5, 100 mM NaCl, 1 mM ZnCl₂. Primer B was prepared as described for the S_1 -nuclease mapping probe except that the synthesized DNA was cut at the *SmaI*(b) site (see *b*). This 232-nucleotide single-stranded primer (800,000 c.p.m., 0.16 pmol) was hybridized with the same MCF-7 RNA as above using the same conditions. After precipitation with ethanol, the primer was elongated with 25 U of reverse transcriptase for 10 min at 42 °C (50 mM Tris-HCl pH 8.3, 40 mM KCl, 8 mM MgCl₂, 0.5 mM dNTP, 0.4 mM dithiothreitol). S_1 -nuclease treatment before gel electrophoresis was as described above. For primer extension mapping with primer A, the primer (5 $\times$ 10⁶ c.p.m., 1 pmol) was hybridized as for primer B, except that the formamide concentration was reduced to 10%. Primer A was elongated with reverse transcriptase as described for primer B, but was not treated with S_1 nuclease. Both the S_1 -nuclease-protected and primer-extended fragments were electrophoresed on a standard 6% sequencing gel.

farther 5' than λ OR8 were obtained when the 5' *SmaI*(b)/*SmaI*(c) fragment of λ OR8 (see Fig. 1) was used as a probe to screen the randomly primed λ gt10 library, which has been described previously¹⁰. Clones λ OR8, λ OR15, λ OR19 and λ OR21 (Fig. 1) were subcloned into the M13 sequencing vector mp8, yielding a series of overlapping fragments suitable for sequencing by the dideoxy technique. The complete sequence is shown in Fig. 2.

The 400-base pair (bp) *SmaI*(b)/*SmaI*(c) fragment of λ OR8 was also used to screen a human leukocyte genomic λ library. The 1.1-kb *PvuII*/*PvuII* fragment of one of the isolated clones (λ GHER1), which contains the *SmaI*(b)/*SmaI*(c) region of λ OR8 together with ~500 bp of 5'-flanking sequence, was subcloned into M13mp18 (Fig. 1). A single-stranded DNA probe, corresponding to the coding strand of the ER gene, was prepared from this *PvuII* fragment using the 5' end-labelled synthetic oligonucleotide primer A (5'-TCAGGGGCTCCAGCTCG-3'; see Figs 2 and 3*b* legends). This probe was used in an S_1 -nuclease mapping assay with sucrose gradient fractions of MCF-7 poly(A)⁺ RNA enriched in ER mRNA¹⁰. The ER mRNA remained homologous to this genomic probe as far as an adenine located 77 bp upstream from the *SmaI*(b) site of λ OR8 (see Figs 2 and 3*b*, arrow +1, and Fig. 3*a*, lane 2, arrowhead; the faint bands which are larger than the major protected fragment may correspond to minor transcription start sites). Primer extension experiments, using two different primers, were performed to show that this site corresponds to the 5' end of the ER mRNA rather than to an intron-exon boundary. Primer A was hybridized to the ER mRNA-enriched RNA and elongated with reverse transcriptase. Primer B was prepared from the *PvuII*/*PvuII* M13 clone, exactly as described above for the S_1 -nuclease mapping experiment, except that the primer was cut at the *SmaI*(b) site so as to generate a 232-nucleotide fragment which was completely homologous to ER mRNA (Fig. 3*a*, lane 5, and *b*). Primer B was hybridized to ER mRNA-enriched RNA and elongated as for primer A (see Fig. 3 legend). After incubation with reverse transcriptase, both primers A (Fig. 3*a*, lane 6, open arrowhead) and B (lane 3, open arrowhead; the band in the control lane 4 may correspond to a self-primed elongation product of the probe) were extended to one of the two guanine nucleotides located at positions +2 and +3 (Fig. 3*a*). We conclude from these results and similar experiments not shown that the A indicated by +1 in Figs 2 and 3 probably corresponds to the 5' end of the ER mRNA. This conclusion is supported by the presence of a TATA box-like sequence, 5'-TACTTAAA-3', located 27 bp upstream from this putative start site (Fig. 2). The sequence 5'-GCCAATGT-3' situated at -103 (Fig. 2) may correspond to a CAT box¹².

The MCF-7 ER mRNA is 6,322 nucleotides long and has a 5'-untranslated region of 232 nucleotides (Figs 1, 2). The most 5' ATG codon, at position +119, is followed 20 codons downstream by a TGA stop codon. A second ATG codon, at position +233, is in the same translational reading frame as the first and is flanked by nucleotides which have a closer homology to the Kozak's consensus sequence¹¹. This codon represents the start of an open reading frame of 1,785 nucleotides which codes for a 595-amino-acid protein of *M*_r 66,182. It is common for translation in eukaryotes to initiate at the most 5' ATG codon of the mRNA¹¹. However, the presence of an upstream ATG codon is not unique to the ER mRNA¹³, although usually only a few codons are present between the ATG and stop codons. It is not known whether this first 20-codon open reading frame is expressed and has any physiological role. In this respect, it will be of interest to compare the human and other ER cDNA sequences to determine whether this region is conserved. Whether the presence of this upstream ATG codon affects the translational efficiency of the ER mRNA is also unknown, but this region does not appear to have the potential to form any stable secondary structure as is observed in some other mRNA sequences¹⁴. A long 3'-untranslated region (4,305 nucleotides)

Fig. 4 Hydropathicity plot of the MCF-7 ER amino-acid sequence and sequence homology with *v-erb-A*. *a*, A computer program, using the algorithm described by Kyte and Doolittle⁴⁰, was used to display the hydropathic nature of the MCF-7 ER using a window of 15 amino acids. Negative values represent hydrophilic regions and positive values hydrophobic regions. The numbers on the horizontal axis refer to the amino-acid sequence beginning at the N-terminus. The region of the ER sequence (residues 185–251) possessing the greatest homology with *v-erb-A* is indicated by a bar.



b, The amino-acid sequences of the MCF-7 ER and *erb-A* were compared by a matrix analysis using a window of six residues and a dot plotted for a match of four or more identical residues. The program therefore generates a diagonal line where any homology between the two sequences is found. The numbers along the axes for both sequences are as described in *a*.

c, The alignment between the sequence of the MCF-7 ER and *v-erb-A* protein was determined using a search procedure based on the amino-acid physical characteristics^{41,42}. There are 378 aligned residues for the ER and *v-erb-A* proteins, of which 27% are identical (boxed amino acids) and 41% conserved according to the following scheme: (P, G), (M, C), (Y, W, F, H), (L, V, I, A), (K, R), (E, Q, N, D) and (S, T) (denoted by a line both above and below the amino acid).

The average correlation coefficient between the two sequences, using six amino-acid physical characteristics, was 0.37. The mean correlation, maximum correlation and standard deviation for control protein sequences (see below) were 0.00, 0.20 and 0.04, respectively. The ER-*v-erb-A* correlation was 9.2 s.d. above the control mean and 4.2 s.d. above the control maximum. The control protein sequence consisted of a group of 63 proteins of unrelated tertiary structure (see ref. 43 for examples) whose sequence was expanded to contain 1,000 residues. Segments of 378 amino acids were randomly selected (using a random generator seeded with a number based on the time of day) and an average correlation coefficient calculated. This procedure was repeated at 10 different times and the mean s.d. and maximum correlations were calculated for the resulting 19,530 average correlations over six amino-acid characteristics. To determine whether the homology between the ER and *v-erb-A* is significant throughout the entire *v-erb-A* sequence, the ER sequence was subdivided into three regions (I, residues 146–255; II, residues 256–467; III, residues 468–595) and each aligned with *v-erb-A* as shown in *c*. The correlation coefficients were 5.8, 6.2 and 3.2 s.d. above the control mean, respectively.

is not unique to the ER; other receptor mRNAs^{15,16}, including that of the human glucocorticoid receptor¹⁷, have a similar structure. The putative polyadenylation signal 'AATAAA' is found 15 bp upstream from the polyadenylation site. Two other such sequences, found in the 3'-untranslated region of the ER mRNA at positions 4,422 and 5,497, are used weakly, if at all, as only the 6.3-kb transcript was detected¹⁰.

ER sequence and homology to *v-erb-A*

The ER amino-acid sequence is shown in Fig. 2 and its hydropathicity is analysed in Fig. 4a. The human ER can apparently be subdivided into at least three hydropathic domains. The first domain, containing the first 120 amino acids, is followed by a region between amino acids 120 and 300 which is rich in cysteine, lysine and arginine and poor in leucine and proline. This hydrophilic region is likely to be exposed on the surface of the receptor and, being rich in positively charged amino acids, may bind to nucleic acids. In this respect, we note that this region may contain the epitope for the monoclonal antibody, H226, which

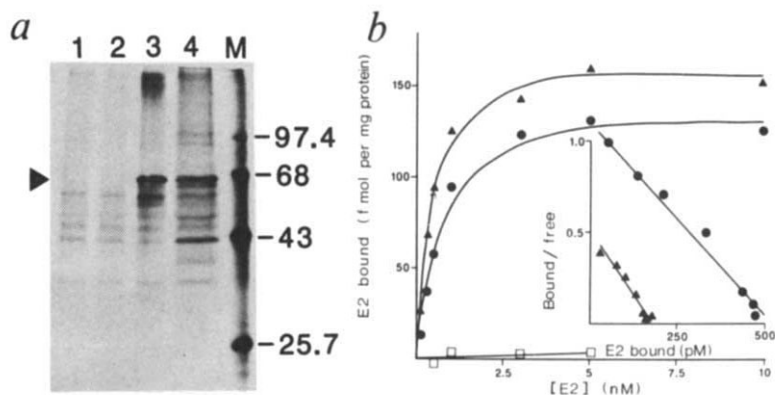
is believed to interact with the DNA-binding domain of the receptor¹⁸ (Fig. 1). A characteristic feature of prokaryotic DNA-binding proteins is the helix- β -turn-helix secondary structure, in which the two helices provide the DNA specificity of the binding through intimate contact with the major groove of the B-form of the DNA helix¹⁹. We compared the ER amino acid sequence with those of the prokaryotic proteins that are either believed or known to contain DNA-binding domains, and with that of the protein thought to be encoded by the *Drosophila* segmentation gene *fushi tarazu* (*ftz*), and the yeast MAT α 1 and MAT α 2 proteins which may contain such domains (ref. 20 and references therein). No structural similarities between any of these proteins and the ER were found, and no similar helix- β -turn-helix secondary structure is predicted from the ER sequence. The third domain corresponds to the C-terminal half of the receptor from approximately amino acid 300 onwards. Being mostly hydrophobic, this domain may contain the oestrogen-binding site.

The ER in MCF-7 cells is predominantly nuclear³. The yeast

Fig. 5 Expression of the ER cDNA in HeLa cells. **a**, The size of the protein coded by the cloned ER cDNA is identical to that of the MCF-7 ER. The ER cDNA sequence was inserted into the eukaryotic expression vector pKCR2 (ref. 29), yielding pKCR2-ER. This recombinant was transfected into HeLa cells and the proteins labelled with ^{35}S -methionine. Cytoplasmic proteins were immunoprecipitated with ER antibodies, separated by SDS-polyacrylamide gel electrophoresis (7%) and revealed by fluorography. Lane 1, wild-type HeLa cells; lane 2, HeLa cells transfected with pKCR2; lane 3, HeLa cells transfected with pKCR2-ER; lane 4, wild-type MCF-7 cells. The arrowhead shows the position of the 65,000- M_r protein. Lane M, ^{14}C -labelled protein markers: phosphorylase b , bovine serum albumin, ovalbumin, α -chymotrypsinogen. Numbers on the right indicate their respective M_r s ($\times 10^{-3}$).

b, The ER cDNA-encoded protein binds oestradiol with the same affinity as the MCF-7 ER. Cytosol extracts, prepared from MCF-7 cells (●) or HeLa cells transfected with pKCR2 (□) or pKCR2-ER (▲), were incubated overnight at 0°C in the presence of ^3H -17 β -oestradiol (E2) with or without a 100-fold excess of unlabelled 17 β -oestradiol. The unbound steroid was removed using dextran-coated charcoal and the receptor-bound steroid counted in a scintillation counter. The number of moles of steroid bound per milligram of cytosolic protein was calculated after subtraction of the nonspecific binding. Each point was performed in duplicate and line-fitting was by computer⁴⁴. The inset shows a Scatchard plot of these data.

Methods. **a**, For construction of vector pKCR2-ER, the ER cDNA clone λOR8 , which had been subcloned into the *Eco*RI site of pBR322, was partially digested with *Sac*II to remove the upstream ATG codon, which is followed by a stop codon (see Fig. 2). The extremities of the DNA were blunt-ended with T4 DNA polymerase and *Eco*RI linkers added. The DNA was digested with *Eco*RI, cutting both in the newly added linker and in the *Eco*RI site present at the 3' end of the pOR8 insert. The 1.8-kb *Eco*RI fragment was then inserted into the *Eco*RI site of pKCR2²⁹. This vector contains the SV40 early promoter and β -globin splicing donor and acceptor sites upstream of the inserted ER sequence, and downstream from it the β -globin and SV40 polyadenylation sites. Transfection of 30–50% confluent HeLa cells (in 9-cm Petri dishes) with calcium phosphate-precipitated supercoiled plasmid DNA (10 μg per dish) was performed as described previously⁴⁵. All the serum used was stripped of endogenous steroid hormones by charcoal treatment. The medium was replaced 48 h after transfection with 3 ml of methionine-free and serum-free Dulbecco's medium containing 100 μCi of ^{35}S -methionine (1,500 Ci mmol⁻¹). After 5 h, the cells were washed with ice-cold phosphate-buffered saline (PBS), resuspended for 5–10 min at 4°C in 200 μl of 1% Nonidet-P40, 50 mM Tris-HCl pH 7.5, and centrifuged at 12,000g for 1 min. SDS (22 μl of 10%) and EDTA (3 μl of 0.2 M) were added to the supernatant and the protein denatured by boiling for 5 min. The ER protein present in HeLa cells (10×10^6 c.p.m. total protein radioactivity) or MCF-7 cells (18×10^6 c.p.m.) was immunoadsorbed to protein A-Sepharose using 5 μg each of the monoclonal antibodies H226 and D75 as described previously¹⁰, before electrophoresis. **b**, HeLa cells were transfected as described in **a**. The cells from 10 Petri dishes were washed twice with ice-cold PBS and homogenized on ice in 1 ml of buffer (20 mM Tris-HCl pH 7.3, 1 mM EDTA, 2 mM β -mercaptoethanol, 50 mM NaCl, 0.3 mM phenylmethylsulphonyl fluoride, 10 mM Na-molybdate and 10% glycerol), using 50 strokes of a 5-ml glass-glass homogenizer. The cytosol was prepared by centrifugation of the homogenate at 105,000g at 4°C for 1 h. Aliquots (50 μl) were incubated in duplicate with different concentrations of ^3H -17 β -oestradiol (153 Ci mmol⁻¹) with (nonspecific binding) or without (total binding) a 100-fold excess of unlabelled 17 β -oestradiol. After incubation overnight at 0°C, 100 μl of dextran-coated charcoal (DCC) solution (0.5% Norit-A, 0.05% dextran T70 in 10 mM Tris-HCl, pH 7.5) was added for 15 min on ice to remove the free steroid. DCC was removed by centrifugation and the amount of ^3H -17 β -oestradiol remaining in 100 μl of supernatant counted. Cytosol preparation and determination of the ER content of MCF-7 cells grown in stripped medium⁹ were carried out as described above for HeLa cells. Cytosolic protein contents were estimated using the Bradford assay.



protein, MAT α 2, contains a sequence of three hydrophobic amino acids (one being proline) with a basic amino acid at each side which may be responsible for the transfer of a MAT α 2- β -galactosidase fusion protein to the nucleus²¹. A similar region (Arg-Pro-Gln-Leu-Lys³²) is present near the N-terminus of the ER (Fig. 2). No region homologous to those sequences responsible for nuclear transfer of the simian virus 40 (SV40) large-T antigen²² or the yeast ribosomal protein L3 (ref. 23) was found.

It has been suggested that tyrosine phosphorylation of the calf uterine ER by a cytoplasmic kinase potentiates the binding of oestrogen²⁴. The selection of the tyrosine phosphorylation site may require the presence of basic (lysine or arginine) and acidic (glutamate or aspartate) amino acids on the N-terminal side of the phosphorylated tyrosine²⁵. Several of the tyrosine residues (positions 43, 184, 219, 526) of the MCF-7 ER sequence exhibit this general feature. Two potential cyclic AMP-dependent phosphorylation sites²⁶ (basic amino acid-basic amino acid-X-Ser/Thr) are present in the ER sequence at positions 236 and 305 (see Fig. 2).

The ER amino-acid sequence was used in a computer search of the NBRF protein database for related sequences. Only one sequence, that of the AEV *gag-erb-A* fusion protein²⁷, showed significantly homology (Fig. 4b, c). This homology is particularly striking in the cysteine-rich region of the ER between amino acids 185 and 251, where each of the 9 cysteine residues in this region is conserved in the *erb-A* protein. This high degree of conservation suggests that the tertiary structure of this region is an important feature of the ER. Although less striking than that within the cysteine-rich region, the homology between the ER and *erb-A* protein extends throughout the *erb-A* sequence (see Fig. 4c). Very little, if any, homology was observed between

the MCF-7 ER sequence and the rat prostatic steroid-binding proteins or rabbit uteroglobin²⁸, which bind weakly androgens and progesterone, respectively (data not shown).

Expression of ER cDNA in HeLa cells

The ER cDNA insert of λOR8 was partially digested with *Sac*II so as to remove the most 5' ATG codon together with the stop codon located 20 codons downstream (Figs 1, 2). The 1.8-kb fragment, containing the complete ER open reading frame, was inserted in the eukaryotic expression vector pKCR2 downstream from the SV40 early promoter²⁹. HeLa cells transfected with this construct (pKCR2-ER) were labelled with ^{35}S -methionine and the proteins immunoprecipitated with the two ER monoclonal antibodies H226 and D75 (ref. 18) (Fig. 5a, lane 3). The endogenous MCF-7 cell ER was similarly labelled and precipitated, and the proteins were separated on a 7% SDS-polyacrylamide gel (Fig. 5a, lane 4). In both cases a protein of $M_r \sim 65,000$ was observed. No such protein was seen in either wild-type HeLa cells (Fig. 5a, lane 1) or HeLa cells transfected with the pKCR2 vector (lane 2). Other smaller proteins, which appear to be specific to the HeLa cells transfected with pKCR2-ER, may be due to either proteolysis of the ER or premature termination of translation.

Cytosol extracts were prepared from either HeLa cells, transfected with pKCR2 or pKCR2-ER, or from MCF-7 cells. A dextran-coated charcoal assay was then used to determine both the number and affinity for oestradiol of any oestrogen-binding protein present in HeLa cells and in MCF-7 (Fig. 5b). HeLa cells transfected with the pKCR2 vector had only background oestradiol-binding activity, whereas HeLa cells transfected with pKCR2-ER contained significant amounts of oestradiol-binding activity (Fig. 5b). The affinity of the corresponding protein for

oestradiol ($K_D = 0.4$ nM) was very similar to that of the MCF-7 ER ($K_D = 0.5$ nM) (Fig. 5b, inset). We conclude that the expression of the cloned ER cDNA in HeLa cells yields a protein which, using these criteria, is indistinguishable from the MCF-7 cell ER.

Discussion

We have isolated and sequenced cDNA clones corresponding to the complete sequence of the human ER mRNA. When expressed in HeLa cells, this cDNA codes for a protein that has the same apparent relative molecular mass and binds oestradiol with the same affinity as the MCF-7 cell ER.

It is extremely interesting that the predicted human ER sequence shares strong sequence homology with the *erb-A* gene of the oncogenic AEV. The AEV genome contains two cell-derived genes termed *v-erb-A* and *v-erb-B* (for review see ref. 30). The *v-erb-A* gene is expressed as a cytoplasmic *gag-erb-A* fusion protein of ~75,000 M_r , which is not itself oncogenic³¹. Instead it appears that its action is to increase the oncogenic potential of *v-erb-B*³¹, which is homologous to the epidermal growth factor (EGF) receptor, although it lacks the EGF-binding domain³². The function of *c-erb-A*, the cellular counterpart of the *v-erb-A* gene, is unknown; it is also unknown whether the *erb-A* protein binds to either DNA or steroid hormones. The chicken ER does not correspond to *c-erb-A*, since its sequence is more closely homologous to that of the human ER than to *v-erb-A* (our unpublished results). The human *c-erb-A* has been mapped to chromosome 17 and part of its amino-acid sequence (residues 226–312) is 96% homologous to *v-erb-A*²⁷; in contrast, the human ER maps¹⁰ to chromosome 6 and shares only a 6.7% homology with *v-erb-A* in this region. Although the non-cysteine-rich domain 2 (ref. 27) of the *v-erb-A* protein shares some homology with human carbonic anhydrase²⁷, no such homology exists between carbonic anhydrase and the human ER (1.3 s.d. above the control mean).

Two *v-erb-A*-related human genomic clones have been isolated³³. The sequences of these two clones are only distantly related³³ and both appear to map³⁴ to human chromosome 17. Southern analysis with human DNA using one of these clones, λ he-A2, suggests the presence of at least an additional related gene showing that *c-erb-A* belongs to a multigene family.

Southern analysis of the human MCF-7 ER gene using λ OR8 as a probe¹⁰ shows that neither λ he-A1 nor λ he-A2³³ corresponds to ER sequence.

It is probable that the conserved cysteine-rich domain, common to the human ER and *v-erb-A* proteins, has some important biological function. It will be interesting to investigate whether it is common to all members of the *erb-A* family. Site-directed *in vitro* mutagenesis of the ER cDNA sequence in the expression vector pKCR2-ER should help to elucidate its function, as well as to localize those regions involved in the multiple functions of the ER. In this respect, a comparison of the ER and *v-erb-A* sequences with that of the glucocorticoid receptor, which has recently been cloned^{17,35,36}, should also be very informative.

There is a strong correlation between the growth of breast cancers and their ER content⁷, and although the mechanism by which the ER increases the growth rate of these tumours is unknown, it may involve the induction of either growth factors or their receptors^{37,38}. The mechanism by which *v-erb-A* potentiates the action of *v-erb-B* is also unknown, but it is possible that the ER may similarly potentiate a human oncogene in breast cancer. In this respect, it will be interesting to investigate whether the sequence of the MCF-7 breast cancer cell line ER differs in some way from that of the ER from non-malignant tissues.

We thank G. L. Greene (University of Chicago) and L. S. Miller (Abbott Laboratories) for providing the D75 and H226 ER monoclonal antibodies, respectively; A. Staub for synthesis of the oligonucleotides; the Michigan Cancer Foundation for providing MCF-7 cells; Transgene for providing the human genomic library; I. Issemann for technical assistance; L. Heydler for growing the cells; and C. Werlé and B. Boulay for the preparation of the figures. This research was supported by INSERM, CNRS (ATP Etats-Unis) and the Ministère de la Recherche et de la Technologie (84VO803). S.G. is a recipient of a Royal Society European Exchange fellowship.

Note added in proof: After this manuscript was submitted, the sequence of the human glucocorticoid receptor and its homology with *v-erb-A* were reported^{46,47}. A comparison of the human and chicken oestrogen receptor, the human glucocorticoid receptor and *v-erb-A* shows that the cysteine-rich and hydrophobic regions are conserved in all four sequences, supporting the view that they correspond to important functional domains⁴⁸.

Received 6 December 1985; accepted 20 January 1986.

- Knowler, J. T. & Beaumont, J. M. *Essays Biochem.* **20**, 1–39 (1985).
- Yamamoto, K. R. *A. Rev. Genet.* **19**, 209–252 (1985).
- King, W. J. & Greene, G. L. *Nature* **307**, 745–747 (1984).
- Welshons, W. V., Lieberman, M. E. & Gorski, J. *Nature* **307**, 747–749 (1984).
- Dean, D. C., Gope, R., Knoll, B. J., Riser, M. E. & O'Malley, B. W. *J. Biol. Chem.* **259**, 9967–9970 (1984).
- Jost, J.-P., Geiser, M. & Seldran, M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 988–991 (1985).
- Edwards, D. P., Chamness, G. C. & McGuire, W. L. *Biochim. biophys. Acta* **560**, 457–486 (1979).
- Lippman, M. E., Bolan, G. & Huff, K. K. *Cancer Res.* **36**, 4595–4601 (1976).
- Masiakowski, P. *et al. Nucleic Acids Res.* **10**, 7895–7903 (1982).
- Walter, P. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 7889–7893 (1985).
- Kozak, M. *Nucleic Acids Res.* **12**, 857–872 (1984).
- Benoist, C., O'Hare, K., Breathnach, R. & Chambon, P. *Nucleic Acids Res.* **8**, 127–142 (1980).
- Hunt, T. *Nature* **316**, 580–581 (1985).
- Dickson, L. A., de Wet, W., Di Liberto, M., Weil, D. & Ramirez, F. *Nucleic Acids Res.* **13**, 3427–3438 (1985).
- Yamamoto, T. *et al. Cell* **39**, 27–38 (1984).
- Schneider, C., Owen, M. J., Banville, D. & Williams, J. G. *Nature* **311**, 675–678 (1984).
- Govindan, M. V., Devic, M., Green, S., Gronemeyer, H. & Chambon, P. *Nucleic Acids Res.* **13**, 8293–8304 (1985).
- Greene, G. L., Sobel, N. B., King, W. J. & Jensen, E. V. *J. Steroid Biochem.* **20**, 51–56 (1984).
- Pabo, C. O. & Sauser, R. T. *A. Rev. Biochem.* **53**, 293–321 (1984).
- Laughon, A. & Scott, M. P. *Nature* **310**, 25–31 (1984).
- Hall, M. N., Hereford, L. & Herskowitz, I. *Cell* **36**, 1057–1065 (1984).
- Kalderon, D., Roberts, B. L., Richardson, W. D. & Smith, A. E. *Cell* **39**, 499–509 (1984).
- Moreland, R. B., Nam, H. G., Hereford, L. M. & Fried, H. M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6561–6565 (1985).
- Mighiaccio, A., Rotondi, A. & Auricchio, F. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5921–5925 (1984).
- Patschinsky, T., Hunter, T., Esch, F. S., Cooper, J. A. & Sefton, B. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 973–977 (1982).
- Krebs, E. G. & Beavo, J. A. *A. Rev. Biochem.* **48**, 923–959 (1979).
- Debuire, B. *et al. Science* **224**, 1456–1459 (1984).
- Baker, M. E. *Biochem. biophys. Res. Commun.* **114**, 325–330 (1983).
- Breathnach, R. & Harris, B. A. *Nucleic Acids Res.* **11**, 7119–7136 (1983).
- Graf, T. & Beug, H. *Cell* **34**, 7–9 (1983).
- Frykberg, L. *et al. Cell* **32**, 227–238 (1983).
- Ullrich, A. *et al. Nature* **309**, 418–425 (1984).
- Jansson, M., Philipson, L. & Vennström, B. *EMBO J.* **2**, 561–565 (1983).
- Spurr, N. K. *et al. EMBO J.* **3**, 159–163 (1984).
- Miesfeld, R. *et al. Nature* **312**, 779–781 (1984).
- Weinberger, C. *et al. Science* **228**, 740–742 (1985).
- Sirbasku, D. A. & Leland, F. E. in *Biochemical Actions of Hormones* Vol. 9 (ed. Litwack, G.) 115–140 (Academic, New York, 1982).
- Mukka, V. R. & Stancel, G. M. *J. Biol. Chem.* **260**, 9820–9824 (1985).
- Frischauf, A.-M., Lehrach, H., Poustka, A. & Murray, N. J. *molec. Biol.* **170**, 827–842 (1983).
- Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105–132 (1982).
- Argos, P., Hanei, M., Wilson, J. M. & Kelley, W. N. *J. Biol. Chem.* **258**, 6450–6457 (1983).
- Argos, P. *EMBO J.* **4**, 1351–1355 (1985).
- Rosmann, M. G. & Argos, P. *A. Rev. Biochem.* **50**, 497–532 (1981).
- Green, S., Field, J. K., Green, C. D. & Beynon, R. J. *B. Nucleic Acids Res.* **10**, 1411–1421 (1982).
- Banerji, J., Rusconi, S. & Schaffner, W. *Cell* **27**, 299–308 (1981).
- Hollenberg, S. M. *et al. Nature* **318**, 635–641 (1985).
- Weinberger, C., Hollenberg, S. M., Rosenfeld, M. G. & Evans, R. M. *Nature* **318**, 670–672 (1985).
- Krust, A. *et al. EMBO J.* (in the press).

Strong breathing of the hydrogen coma of comet Halley

E. Kaneda*, K. Hirao†, M. Takagi‡, O. Ashihara§,
T. Itoh§ & M. Shimizu§

* Geophysics Research Laboratory, University of Tokyo, Hongo,
Bunkyo-ku, Tokyo 113, Japan

† Department of Aeronautics and Astronautics,
Tokai University, 1117, Kita-kaname, Hiratsuka-shi,
Kanagawa Prefecture 259-12, Japan

‡ Institute of Industrial Science, University of Tokyo, Roppongi,
Minato-ku, Tokyo 106, Japan

§ Institute of Space and Astronautical Science, Komaba, Meguro-ku,
Tokyo 153, Japan

Strong breathing of the hydrogen coma of comet Halley was observed by the Lyman- α imager of the Japanese interplanetary spacecraft, Suisei. Measurement of the peak-to-peak variation of central brightness yields a preliminary value of 2.2 ± 0.1 days for the period. Two outbursts from the active regions on the comet's surface could be the cause of this variation. A shell structure in the hydrogen coma was also detected. The outbursts appear to contribute overwhelmingly to the Lyman- α activity of this comet, at a pre-perihelion heliocentric distance of ~ 1.5 AU.

Comet Halley has a reddish dark nucleus with a radius of 3 ± 1 km (refs 1, 2). Various observations of OH, H and other volatiles suggest that the core beneath this dark mantle could be the usual dirty ice. Water, the main constituent of the coma, is photodissociated into OH and H by solar ultraviolet radiation, resulting in a high-velocity component (~ 20 km s $^{-1}$) of the hydrogen coma. Further photodissociation of OH creates a low-velocity component (~ 8 km s $^{-1}$). In addition to these major processes, there are many minor photochemical processes which produce hydrogen atoms with various velocities³.

The analysis of the dust behaviour of comet Halley in the 1910 apparition^{4,5} suggested that water production was not necessarily uniform, a significant contribution arising from outbursts (with velocities of several hundred metres per second) at three locations on the surface (two point sources and one crack at almost the same longitude). It appears that an outburst occurs when, due to the rotation of the comet nucleus, one of these sources faces the Sun. If this assumption is correct, it can be expected that analysis of the intensity variation of the hydrogen coma will determine the rotational period of the comet.

The Lyman- α (1216Å) imager onboard Suisei, the first Japanese interplanetary spacecraft, is a charge-coupled device (CCD) imaging detector with a field of view of $2.0^\circ \times 1.8^\circ$ (153×122 pixels) (ref. 6 and E.K., K.H. and M.T., in preparation). This imager took several hundred faint ultraviolet images of comet Halley during the 3 weeks following 26 November 1985 (see Fig. 1). Figure 2 shows the synthesized images of Halley at 00:00 30 November UT (active phase) and 04:00 28 November UT (inactive phase), respectively, obtained by superposing seven images on its first position, after adjusting for the comet's subsequent motion.

The time-sequential analysis of the images showed the following activity of comet Halley (see also Fig. 1): During an inactive phase (Fig. 2b), the images were quite faint (04:00 28 November). After that phase, an expansion of the hydrogen coma was observed, possibly arising from the photodissociation of a weak outburst. About 20 h after the blasting of this weak outburst, a bright pixel (which corresponds to a linear size of about 40,000 km) appeared suddenly adjacent to the ephemeris position of the comet. The transit time over this distance of a hydrogen atom with a velocity of 10 km s $^{-1}$ is about 1 h; thus,

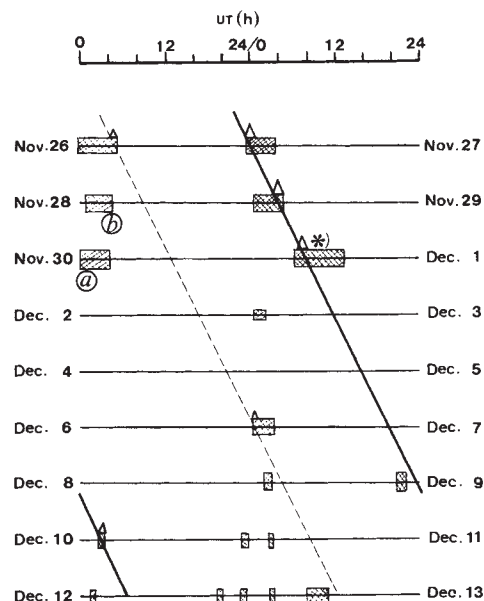


Fig. 1 Time-sequential plot of observation intervals and activities of the hydrogen coma. Cross-hatched sections represent the observational time intervals. Triangular spikes indicate the times when the images taken included central bright pixels, suggesting the appearance of a jet. Diagonal lines crossing the figure have a slope of 2.2 days periodicity (solid line, strong outburst; broken line, weak outburst). The asterisk indicates that the images on this interval are taken in a compressed mode (to save data transmission time). The times when the images in Figs 2 and 3 were taken are indicated (@, ⑥).

the phenomenon could be interpreted as the occurrence of a strong outburst. These events are observed repeatedly at 23:00 26 November, 04:00 29 November and 09:00 1 December. With the appearance of this new hydrogen atom source, the images become brighter (Fig. 2a), then gradually fade away as the hydrogen atom density decreases with the expansion of the hydrogen coma. The ultimate image is similar to Fig. 2b. A very similar sequence of events was observed during the week starting 9 December. In particular, the flash at 04:00 10 December was essential in determining the period of the Lyman- α radiance variation.

The difference between the active and inactive phases of the hydrogen coma suggests that the outburst phenomena overwhelm the other activity of this comet, at a heliocentric distance of ~ 1.5 AU. The time interval between flashes (strong outbursts) suggests a preliminary value of 2.2 ± 0.1 days for the period of this activity. The rotational period of comet Halley is still controversial, but the above value is close to 2.23 days, the value derived by Sekanina and Larson^{4,5} from the analysis of dust behaviour in the 1910 apparition. It is also about five times the period of 10 h estimated by Whipple⁷, and about twice the value proposed by Le Fevre *et al.*⁸. This last value has been criticized by West and Pedersen⁹.

Expansion of the hydrogen atoms produced from photodissociation of the water outburst will form a shell structure, propagating with time in the hydrogen coma, because the velocity of the hydrogen atoms is much higher than that of the parent molecules. The image in Fig. 2a shows such a shell, with a radius of about 0.58×10^6 km. At the comet's present location, the photodissociation time of water is about 2 days, comparable to the characteristic travel time of the hydrogen atoms in the coma; therefore, the width of the shell is very large. When Suisei

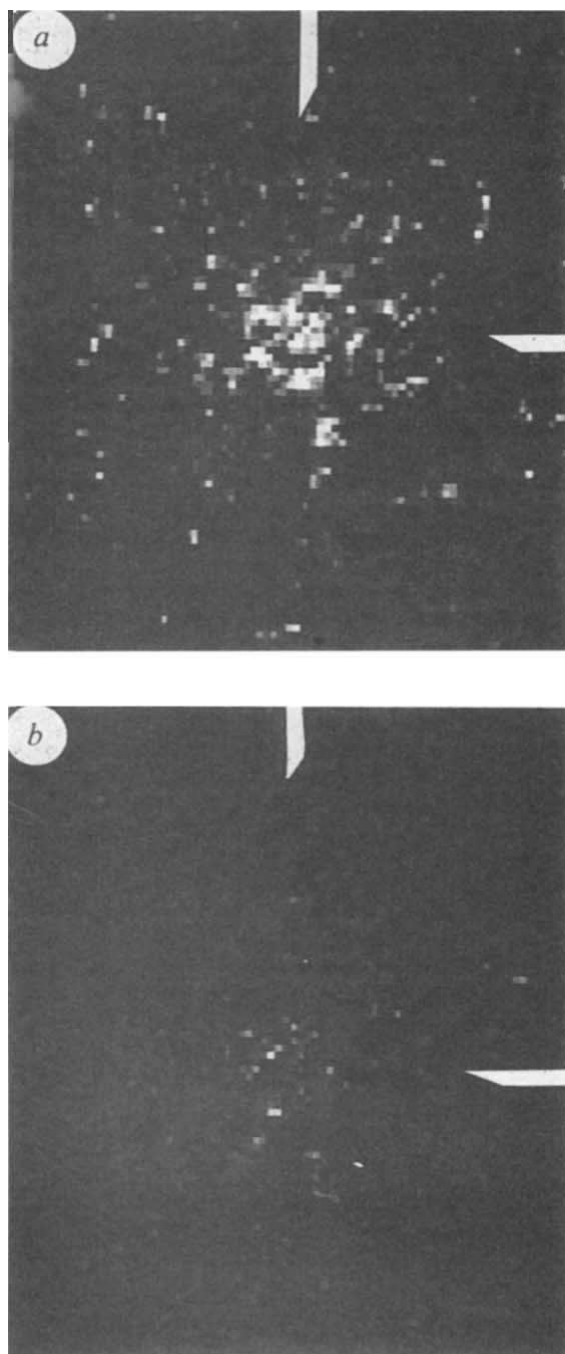


Fig. 2 Lyman- α images of the hydrogen coma of comet Halley taken at *a*, 00:00 30 November UT (active phase) and *b*, 04:00 28 November UT (inactive phase). The vertical and horizontal dimensions of the photographs are both 2.5×10^6 km (at the comet's position). The comet nucleus (located at the intersection of the white marker wedges) is at (right ascension, declination) = (01 h 41 m, $15^\circ 50'$) on the 1950.0 star atlas in *a* and (01 h 56 m, $16^\circ 50'$) in *b*. The distances of comet Halley from the Sun and Suisei are *a*, 1.50 and 0.84 AU, and *b*, 1.53 and 0.83 AU. The faintness of this image suggests that the outburst phenomena overwhelm the other activity of this comet.

encounters the comet, on 8 March 1986, we would expect the shell to have a more sharply defined form due to the shorter photodissociation time. Our continuous observations over the next several months will allow detailed spatial and temporal studies of the hydrogen coma in comet Halley.

We thank the many people in ISAS engaged in the Planet-A mission project; in particular, Professor M. Oda for valuable discussions; also Dr J. Hughes (Harvard-Smithsonian Center for Astrophysics) for his linguistic help.

Received 31 December 1985; accepted 23 January 1986.

1. Wyckoff, S. *et al.* *Nature* **316**, 241-242 (1985).
2. Wallis, M. K. & Wickramasinghe, N. C. *Mon. Not. R. astr. Soc.* **216**, 453-458 (1985).
3. Keller, H. U. *Space Sci. Rev.* **18**, 641-684 (1976).
4. Sekanina, Z. & Larson, S. M. *Astr. J.* **89**, 1408-1425 (1984).
5. Sekanina, Z. *Proc. Interagency Consulting Group, Working Group 1*, 5 (IACG, Pasadena, 1985).
6. Hirao, K. & Kaneda, E. *Adv. Space Res.* **2**, 167-169 (1983).
7. Whipple, F. L. in *Comets* (ed. Wilkening, L. L.) 227-250 (University of Arizona Press, 1982).
8. Le Fevre, O. *et al.* *Astr. Astrophys.* **138**, L1-L4 (1984).
9. West, R. M. & Pedersen, H. *Astr. Astrophys.* **138**, L9-L10 (1984).

Electromagnetic angular momentum transport in Saturn's rings

C. K. Goertz*†, G. E. Morfill†, W. Ip‡, E. Grün§ & O. Havnes||

* Department of Physics and Astronomy, University of Iowa, Iowa City, Iowa 52242, USA

† Max-Planck-Institut für Physik und Astrophysik, Institut für extraterrestrische Physik, 8046 Garching, FRG

‡ Max-Planck-Institut für Aeronomie, D-3411 Katlenburg-Lindau, FRG

§ Max-Planck-Institut für Kernphysik, D-6900 Heidelberg, FRG

|| Auroral Observatory, University of Tromsø, P.B. 953, N-9001, Tromsø, Norway

The observed 'spokes' in Saturn's rings have been interpreted as consisting of elevated, sub-micrometre sized dust particles¹. Arguments in favour of this interpretation are, for example, the photometric properties (spokes are dark in backscattered, bright in forward scattered light), the dynamics (approximate keplerian rotation) and lifetime (less than half an orbital period). We show here that submicrometre dust particles sporadically elevated above the ring are subject to electromagnetic forces which will reduce their angular momentum inside synchronous orbit and increase it outside. When the dust is reabsorbed by the ring the angular momentum of the ring is decreased (increased) inside (outside) of synchronous orbit. For the case of the spokes in Saturn's B-ring we estimate that the timescale for transporting ring material due to this angular momentum coupling effect is comparable to the viscous transport time or even smaller. We suggest that the minimum in the optical depth of the B-ring at synchronous orbit is due to this effect.

The temporal evolution of Saturn's rings is believed to be primarily determined by the angular momentum transport as a consequence of collisions between the differentially rotating ring particles². In addition, mass erosion effects, meteor impacts, resonances with satellites and moonlets and density waves³⁻⁵ are important. So far, electromagnetic effects have not been considered to be important because the electromagnetic forces on the typical big (~ 1 m) ring particles are small compared with the gravitational force of the planet⁶. Only submicrometre dust particles are significantly influenced by electromagnetic effects. We now discuss a potentially very important indirect mechanism whereby electromagnetic forces can affect the angular momentum transport in the ring and hence its long-term evolution. This indirect mechanism works through the small dust particles constituting the spokes. Note that if the spokes do not consist of elevated dust, but only of a rearrangement of dust particles in the ring plane, our mechanism would probably not be relevant. However, as mentioned earlier, the available evidence favours elevated dust, and then it is easily demonstrated that the effect is important. The spokes of Saturn have an optical depth of

$\tau \approx 0.1$, and cover a fraction $\eta \approx 0.05$ of the B-ring between $1.7R_p$ and $1.9R_p$ (planetary radii). The mean particle size is $s = 3 \times 10^{-5}$ cm. The total mass in the spokes is thus

$$M_s = 0.7 \frac{4\pi}{3} s \tau \rho_s \eta R_p^2$$

For a nominal material density of the spoke particles $\rho_s = 1 \text{ g cm}^{-3}$ we obtain $M_s \approx 3 \times 10^{13} \text{ g}$ (see also ref. 6). We argue below that during the time a spoke particle is elevated it has an average radial velocity of $\sim 10^2 \text{ cm s}^{-1}$. Thus the total mass of the B-ring ($7 \times 10^{21} \text{ g}$) could be transported over a distance of $0.2 R_p$ in 10^8 yr .

Consider a dust particle of mass m ejected at a radial distance r out of the ring plane. It has a charge Q . Its velocity is $v_r \hat{r} + v_\phi \hat{\phi}$. If the planetary magnetic field corotates, the particle will be subject to a Lorentz force

$$\vec{F} = m \frac{d\vec{v}}{dt} = Q[v_r \hat{r} + (v_\phi - \Omega_p r) \hat{\phi}] \times \frac{\vec{B}}{c} - \frac{GM_p m}{r^2} \hat{r} \quad (1)$$

where Ω_p is the rotation rate of the planet, G the gravitational constant and c the speed of light. From the ϕ component of equation (1), we get (ϕ is in the direction of the ring particle rotation)

$$\frac{d}{dt} \mathcal{L} = v_r \frac{a}{r^2} \quad (2)$$

where $\mathcal{L} = r v_\phi$ is the specific angular momentum of the particle. We have assumed a dipole field, near the equatorial plane this is

$$\vec{B} = -B_0 \left(\frac{R_p^3}{r^3} \right) \hat{z} \quad (3)$$

and have introduced

$$a \equiv \left(\frac{QB_0}{mc} \right) R_p^3$$

For Saturn $B_0 = 0.2 \text{ G}$, and the planetary radius is $R_p = 6 \times 10^9 \text{ cm}$. When the particle is reabsorbed by the ring after a mean time, t_L , it has changed its specific angular momentum by an amount

$$\Delta \mathcal{L} = \int_0^{t_L} \frac{d\mathcal{L}}{dt} dt \quad (4)$$

Since we are concerned here with particle populations, it is convenient to introduce the angular momentum per unit surface area

$$\tilde{\mathcal{L}} \equiv \sigma \mathcal{L} \quad (5)$$

where σ is the surface density. In the following, we shall use subscripts R and S for rings and spokes, respectively. The rate of change of ring angular momentum can then be easily calculated.

Since in the absorption process of spoke material the total angular momentum must be conserved, the only angular momentum lost by the system is that calculated from equations (2)–(5). Hence,

$$\frac{d\tilde{\mathcal{L}}_R}{dt} = \frac{d\tilde{\mathcal{L}}_S}{dt} \quad (6)$$

The rate of change of the specific ring angular momentum is then

$$\frac{d\mathcal{L}_R}{dt} = \frac{\sigma_S v_r a}{\sigma_R r^2} \quad (7)$$

from which we may compute the timescale for angular momentum change of the ring

$$\tau_R \equiv \mathcal{L}_R \left(\frac{d\mathcal{L}_R}{dt} \right)^{-1} = \frac{\sigma_R r^2 \mathcal{L}_R}{\sigma_S v_r a} \quad (8)$$

For all practical purposes, we may put $\mathcal{L}_R \equiv \mathcal{L}_K$, the Kepler

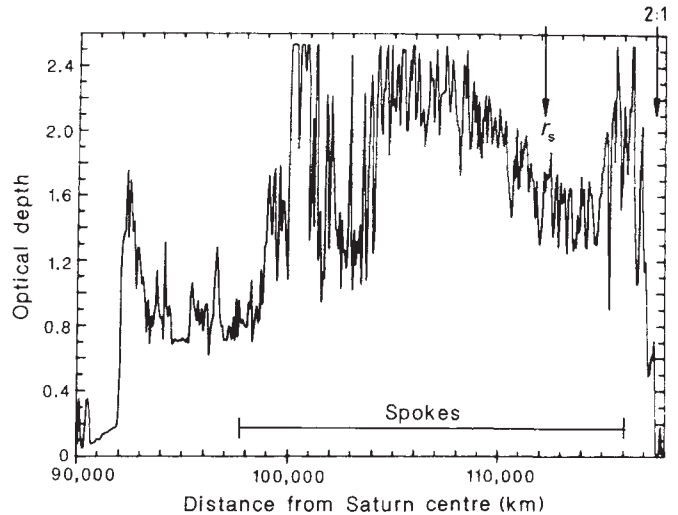


Fig. 1 Optical depth variation in Saturn's B-ring¹⁰. The horizontal line indicates the observed extent of spoke activity. The synchronous radius and the Mimas 2:1 resonance positions are also shown.

value $\sqrt{GM_p/r}$. The radial evolution of the ring material is characterized by a velocity

$$w \equiv \frac{2r}{\mathcal{L}_R} \frac{d\mathcal{L}_R}{dt} = \frac{2\sigma_S v_r a}{\sigma_R r \mathcal{L}_R} \quad (9)$$

The radial velocity at impact can be calculated from equation (1).

$$v_r \approx - \left(\frac{QB_0}{mc} \right) \frac{R_p^3}{r^2} \left(\frac{v_\phi}{r} - \Omega_p \right) t_L \quad (10)$$

To simplify matters we assume that the azimuthal velocity of the particle is everywhere the keplerian velocity. This is a good approximation as long as $\Delta r/r \ll 1$. Thus the effective radial mass transport velocity is (we use $L \equiv r/R_p$)

$$w \approx - \left(\frac{QB_0}{mc} \right)^2 \frac{R_p}{L_S^5} \frac{\sigma_S}{\sigma_R} \left[\left(\frac{L_S}{L} \right)^5 - \left(\frac{L_S}{L} \right)^{7/2} \right] t_L \quad (11)$$

Inside the synchronous orbit ($r_s = R_p L_S$, where $\Omega_k = \Omega_p$) the mass transport is inwards, it is zero at synchronous orbit and outwards beyond r_s .

We note that the direction of the mass transport is independent of the charge sign. The reason is that inside synchronous orbit keplerian dust particles move faster than the corotating magnetic field and, therefore, always lose angular momentum to the planet (irrespective of the charge sign) whereas they gain angular momentum from the planet outside synchronous orbit. This change of angular momentum of one component determines the evolution of the whole ring material. If the small particles remained in the ring they would quickly collide with, and be absorbed by, the large ring particles. In that case t_L would be small. If the elevated particles were bigger $\sigma_S = 4/3 \rho_S s \tau$ increases. However, if the charge is proportional to particle radius, s , (implying constant potential), w is inversely proportional to m , and the significance of the effect is reduced accordingly. The angular momentum coupling to the planet is thus most significant if small dust particles can be elevated above the ring plane, that is, where spoke-like activity occurs in the ring. These considerations suggest that the spokes in Saturn's B-ring, so far regarded as a curious phenomenon without major consequences, may significantly affect the long term evolution of the B-ring itself.

To make some numerical estimates of the effect, we consider the following. The spokes in Saturn's B-ring presumably consist of ice grains with an average size of $s = 3 \times 10^{-5} \text{ cm}$ ($m = 10^{-13} \text{ g}$).

The azimuthally-averaged mass density is $\sigma_s = \tau_s \eta m / \pi s^2 = 1.5 \times 10^{-6} \text{ g cm}^{-2}$ for $\tau_s = 0.1$ and the fractional area covered, $\eta = 0.05$ (ref. 6). If spokes exist below the Voyager contrast or resolution level, which seems likely, then our estimates can be regarded as a conservative lower limit. The lifetime of a spoke is $t_L = 9 \times 10^3 \text{ s}$ (ref. 7). The average charge of a spoke particle is not well known. Goertz⁸ argues that $Q = -200 e$ whereas Eplee and Smith⁹ suggest that Q can be as large as $Q = -10,000 e$. Here e is the electronic charge. The ring surface density σ_R is taken to be 60 g cm^{-2} (ref. 2). The effective radial mass transport velocity is

$$w \approx -1.2 \times 10^{-10} \left(\frac{Q}{e} \right)^2 \left[\left(\frac{L_s}{L} \right)^5 - \left(\frac{L_s}{L} \right)^{7/2} \right] \text{ cm s}^{-1}$$

At $L = 1.8$ the velocity w is $7.0 \times 10^{-12} (Q/e)^2 \text{ cm s}^{-1}$. The transport velocity due to the ring viscosity is given by $-\nu/r$ with $\nu \approx 20 \text{ cm}^2 \text{ s}^{-1}$ a typical value in the B-ring¹. At $L = 1.8$ this is $\sim 2 \times 10^{-9} \text{ cm s}^{-1}$. Thus mass transport due to electromagnetic angular momentum coupling to Saturn is of comparable importance even for $Q/e = 30$ except very close to the synchronous orbit ($L = L_s = 1.86$). We note that the radial velocity of the spoke particles $v_r = -4.3 (Q/e) \text{ cm s}^{-1}$ at $L = 1.8$. The electromagnetic angular momentum timescale, as defined by equation (8), becomes

$$\tau_R = 2.2 \times 10^{18} \left(\frac{Q}{e} \right)^{-2} L^6 \left[1 - \left(\frac{L}{L_s} \right)^{3/2} \right]^{-1} \text{ s}$$

At $L = 1.8$, and for $Q/e = 100$, this yields about $5 \times 10^9 \text{ yr}$. Viscous time scales, $\tau_{\text{vis}} \equiv r^2/\nu$, on the other hand give $\sim 2 \times 10^{11} \text{ yr}$.

The viscosity of a particulate disk is a complex and not well understood function of the ring mass density². In general, there will always be some particle diffusion due to collisions, as well as the systematic radial drift $-\nu/r$. We will discuss this elsewhere. Here we just note that a steady-state solution of the transport problem, including viscous and electromagnetic effects, will display a minimum in the B-ring near, but outside of the synchronous orbit. Figure 1 shows the optical depth variation of the B-ring on a linear scale. It clearly shows such a minimum. If the optical depth of the B-ring was originally constant at -2.5 , corresponding to the values outside the 'synchronous minimum', then the equivalent displaced mass corresponds to $9 \times 10^{20} \text{ g}$, about 13% of the current ring mass. The timescale for displacing this mass can be estimated from the above considerations, and is significantly less than the lifetime of Saturn. If the electromagnetic angular momentum transport process really works as efficiently as we have calculated (this depends on the charge, mass and elevation of spoke particles) this implies that Saturn's ring system is not primordial, or that spoke activity has not been continuous.

We have shown that the sporadic elevation of small dust particles above Saturn's ring, and their subsequent reabsorption, provides a mechanism for angular momentum coupling to the planet, provided the dust particles are charged. There is clear evidence that the spoke particles are, indeed, charged. Then the ring loses angular momentum inside synchronous orbit and thus moves towards the planet, whereas it gains angular momentum and moves away from the planet outside synchronous orbit. In this picture, the outer boundary of the B-ring follows quite naturally.

At $L = 1.9485$, angular momentum is transferred by strong gravitational torques to the moon Mimas (2:1 resonance). In this way, excess angular momentum is removed, and the ring material is clamped in place. The positional agreement between the outer B-ring boundary and the Mimas 2:1 resonance is very good.

One may even speculate that the A-ring has evolved away from this resonance under the past influence of spoke-like

phenomena, or even due to the current action of weak activity below the Voyager observational threshold.

Using reasonable values for the charge of the spoke particles, and observed ring and spoke properties, we have shown that the timescales for B-ring evolution due to the momentum coupling discussed here are comparable with, or even smaller than, viscous evolution timescales.

This effect, combined with some stochastic spreading, should produce a minimum in the optical depth of the ring near, but just outside synchronous orbit. Such a minimum seems, indeed, to exist. We also note that the transport due to the momentum coupling should affect the viscous instability discussed in connection with the formation of ringlets² as well as the resonantly driven density and bending waves. Since spokes are observed only in the outer B-ring this may be less important for the waves in the A and C rings.

In addition, the momentum coupling introduced across the ring by the processes discussed here, may lead to instabilities by themselves. This will be discussed elsewhere (C.K.G. and G.E.M., in preparation).

C.K.G. thanks the NSF for support under grant ATM 81-11126 and NASA contracts NSC-7632, NGL-16-001-043 and NAGW-364. The hospitality of the Auroral Observatory, University Tromsø, in organizing a 'midnight sun' workshop on 'dust plasmas' is gratefully acknowledged.

Received 5 August 1985; accepted 10 January 1986.

1. Terrile, R. J., Yagi, G., Cook, A. F. & Porco, C. C. *Bull. Am. astr. Soc.* **13**, 728 (1981).
2. Stewart, G. R., Lin, D. N. C. & Bodenheimer, P. in *Planetary Rings* (eds Greenberg, R. & Brahic, A.) 447 (University of Arizona Press, Tucson, 1984).
3. Durisen, R. M. in *Planetary Rings* (eds Greenberg, R. & Brahic, A.) 416 (University of Arizona Press, Tucson, 1984).
4. Shu, F. H. in *Planetary Rings* (eds Greenberg, R. & Brahic, A.) 413 (University of Arizona Press, Tucson, 1984).
5. Dermott, S. in *Planetary Rings* (eds Greenberg, R. & Brahic, A.) 589 (University of Arizona Press, Tucson, 1984).
6. Grün, E., Morfill, G. E. & Mendis, D. A. in *Planetary Rings* (eds Greenberg, R. & Brahic, A.) 275 (University of Arizona Press, Tucson, 1984).
7. Grün, E., Garneau, G. W., Terrile, R. J., Johnson, T. V. & Morfill, G. *Adv. Space Res.* **4**, 143-xxx (1984).
8. Goertz, C. K., *Adv. Space Res.* **4**, 137-xxx (1984).
9. Eplee, E. & Smith, B. *Icarus* **63**, 304 (1985).
10. Esposito, L. W. *et al. J. geophys. Res.* **88**, 8643 (1983).

Plasmatization and recondensation of the saturnian rings

W.-H. Ip

Max-Planck-Institut für Aeronomie, D-3411 Katlenburg-Lindau, FRG

One major puzzle from the ground-based and Voyager observations of the saturnian system concerns the presence of the E ring and its possible relationship to the icy satellite, Enceladus¹⁻³. Several issues concerning the surface property and interior condition of Enceladus remain unresolved if it is the main supplier of the E-ring particulate matter. Here we explore the intriguing alternative that the mass supply and orbital configuration of the E ring may be derived from a mass and momentum coupling between the A ring and the tenuous E ring a process of plasmatization and recondensation of the A-ring icy material. In the same way, the C ring could be replenished by the B ring. In this new view, plasma transport has a very important role in the dynamical evolution of the saturnian ring system.

The possible importance of plasma transport in the redistribution of the mass of Saturn's rings has been discussed elsewhere⁴⁻⁸. The basic idea is that charged particles situated at the ring plane would be lost to the planetary surface by field-aligned motion if the radial distance is smaller than a certain critical value. For charged particles initially in co-rotation, for which the magnetic moment is negligibly small, the critical radius is $R_c = 1.625 R_s$ (where R_s is the radius of Saturn); and for

keplerian-launched charged particles with their magnetic moments defined by the difference between the co-rotational speed and the keplerian speed, the critical radius is $R_c^* = 1.525R_s$, as determined by Northrop and Hill⁵ using the Z3 magnetic model⁹ and the J_2 and J_4 terms¹⁰ of the gravitational field of Saturn.

The critical radius R_c is located near a sharp transition in brightness in the B ring at $R \approx 1.64R_s$, and the second critical radius R_c^* is situated only 20 km away from the inner edge of the B ring⁴⁻⁷. Plasma transport could therefore be instrumental in defining the ring mass distribution. Note that although Northrop and Hill considered the relevant charged particles to be tiny icy grains with a very large value of q/m (where q and m are, respectively, the charge and mass of the particles), I stressed the role of large water cluster ions in the mass-loss process (or siphon flow)⁶. In principle, these two types of charged particle are the same as long as the charge-to-mass ratio is effectively very large. In the following we refer to the charged particles under discussion as water cluster ions and their parent neutral particles as water clusters, to distinguish their ring origin. If the loss rate of the water cluster ions remains the same throughout the existence of the ring system, which may be as short as 10^8 yr from dynamical considerations¹¹⁻¹⁴, the average mass injection rate from the ring plane for $R < 1.63R_s$ to the planetary atmosphere would be $\dot{\sigma} \approx \sigma/10^8 \text{ yr} \approx 3.2 \times 10^{-14} \text{ g cm}^{-2} \text{ s}^{-1}$ or $\dot{M} \approx 4 \times 10^6 \text{ g s}^{-1}$ if the surface density is $\sigma \approx 100 \text{ g cm}^{-2}$. This large influx of water ions and molecules could be very important in modifying the ionosphere of Saturn (by electron depletion), as discussed elsewhere^{6,15,16}.

One consequence of the ionization of the ring material is that the water cluster ions will be accelerated or decelerated to the co-rotational motion depending on where they are produced. For those stably trapped in field-aligned motion (that is, not lost to the planetary atmosphere because of the siphon flow effect), they may be directly reabsorbed by the rings during their bounces across the ring plane. As the water clusters recondense on the surfaces of larger ring particles, they will be decelerated or accelerated to keplerian motion. This implies a very efficient method of exchange of angular momentum between the rings and Saturn. For orbital distances larger than the co-rotation distance ($R_{\text{syn}} = 1.86R_s$), there will be an acceleration effect with angular momentum transfer from Saturn to the rings, whereas the opposite effect holds for $R < R_{\text{syn}}$ (ref. 8). It can be estimated that the ring torque from ionization and recycling of the ring material is of the order of $\dot{L} \sim 5 \times 10^{-21} \text{ g cm}^2 \text{ s}^{-1}$ as far as the A ring is concerned¹⁷; this value is comparable to the torque in the opposite direction from gravitational interaction with the shepherding satellites and the co-orbital satellites^{11,18}. It is thus clear that the electrodynamic effect from ring plasmatization could be of importance in holding the A ring in place.

In addition to direct recondensation of the water (ion) clusters, another way to transfer mass and angular momentum is through accretion of the particles flung into elliptical keplerian orbits or ballistic escape orbits after their neutralization from the ring plasma. For cold ring plasma (as indicated by the brightness variation at $1.63R_s$), the keplerian orbit of a neutralized particle will be determined by the radial distance (R_0) at neutralization and the angular momentum (L_0). From conservation of angular momentum and the condition of force balance of a particle in circular keplerian motion, the radial position of this 'projected' particle (R_k) can then be determined.

Because the planet itself represents a perfect absorber, particles with flight paths intercepting Saturn will be lost after the first launch. This condition then defines the inner limit of the inward projection of the plasmatized ring matter (originating from $R < R_{\text{syn}}$). Our computation, taking into account the J_2 and J_4 terms of the gravitational field, indicates that this limit is given as $R_{k,\text{min}} = 1.258R_s$ for $R_{0,\text{min}} = 1.691R_s$. If the C ring forms by the accretion of ring material injected from the B ring to balance the loss through siphon flow, the consideration of

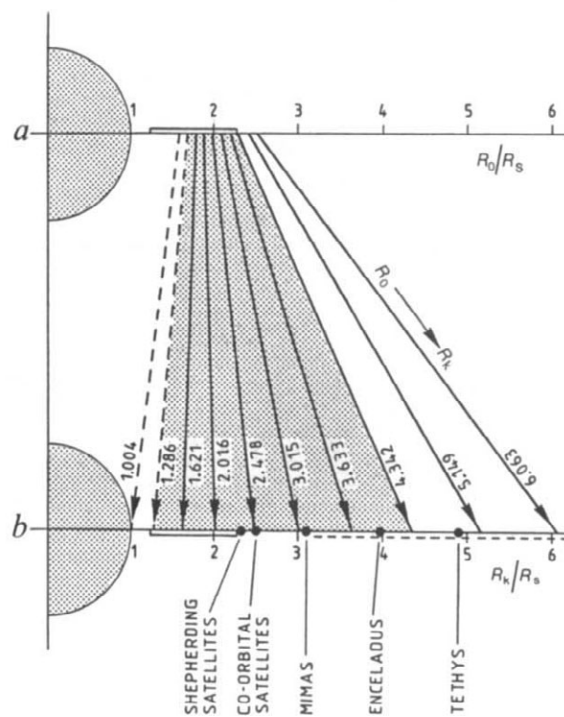


Fig. 1 Angular momentum projection of the ionized ring material in the form of charged particles with $q/m \rightarrow \infty$ after their neutralization. Conservation of angular momentum allows the determination of the orbital radius (R_k) of a particle flung into keplerian orbit (b) from its radial distance (R_0) at neutralization (a). Because of the absorption effect of the planet, neutral particles ejected inside $R = 1.69R_s$ from co-rotating motion will be lost after the first launch. This process may define the inner edge of the C ring. At the same time, material ejected from the A ring would define the spatial extension of the E ring via angular momentum redistribution.

the angular momentum budget would imply that the inner boundary of the C ring should be given approximately by $R_{k,\text{min}}$. The value of $R_{k,\text{min}}$ is indeed almost the same as the observed inner edge of the C ring taken to be at $1.235R_s$. The slight discrepancy could be explained, for example, by the thermal spread of the water cluster ions or collisional scattering effects.

The E ring, which has a peak brightness near the orbit of Enceladus, has generally been thought to be the product of mass ejection from the surface of this icy satellite¹⁻³. The E-ring particles, with an average size determined to be a few micrometres¹⁹, have very short lifetimes against energetic charged particle sputtering and plasma drag (outward) transport^{20,21}. Both mechanisms typically have timescales of the order of 10^3 – 10^4 yr. If the E-ring system is to be maintained as a permanent structure, it will require a steady source²². Below, we consider the possible contribution from the material ejected from the main ring system.

As shown in Fig. 1, the particles neutralized from the ring plasma at different locations in the A ring will have angular momentum corresponding to R_k between $3.0R_s$ and $4.0R_s$. This means that interception of these particles in ballistic flight by the E ring beyond $4R_s$ will lead to braking of the outward motion of the charged E-ring particles as a result of interaction with co-rotating thermal plasma. Furthermore, accretion of these water clusters will lead to replenishment of the material lost by the surface sputtering effect. The mass erosion rate of the E ring may be estimated as follows. For an average optical depth of

$\tau_E \approx 3 \times 10^{-7}$, a total area of $S_E \approx 1.7 \times 10^{21} \text{ cm}^2$ and an average particle size of $d \approx 3 \mu\text{m}$, the total mass erosion rate can be given as $\dot{M}_E \approx -0.6 \text{ g s}^{-1}$ if the sputtering timescale is 10^4 yr .

The injection rate and accretion rate of the water clusters from the A ring are more difficult to derive because the whole sequence of ionization, neutralization and final collisional assimilation must be very complicated. But we could use the C ring as a reference point in the sense that the maintenance of the C ring is conditioned by balancing the mass loss rate (through siphon flow) with the mass accretion rate (from the B ring):

$$\frac{\tau_c S_c}{t_s} = \frac{\tau_B S_B}{t_a} \quad (1)$$

where t_s is the loss timescale and t_a the accretion timescale. From equation (1), $t_a \approx (\tau_B/\tau_c)t_s$ or $t_a \approx 20t_s$ (for equal areas: $S_c \approx S_B$) as the optical depth of the C ring is $\tau_c \approx 0.1$ and that of the B ring $\tau_B \approx 2$. When applied to the E ring, the effective value of the accretion timescale of matter injected from the A ring should be scaled by the probability of interception (P). As $P_c \approx 0(1)$ and $P_E \approx 0(10^{-5})$, the total mass accretion rate will be

$$\dot{M}_E^* \sim \frac{10^{-5} f \tau_A S_A}{t_a} \quad (2)$$

where $f = 100 \text{ g cm}^{-2}$ is the factor converting optical depth to surface density, $\tau_A \approx 0.5$, and $S_A \approx 10^{20} \text{ cm}^2$. Thus, for $t_a \approx 2 \times 10^9 \text{ yr}$, $\dot{M}_E^* \approx 0.8 \text{ g s}^{-1}$; and it can be seen that $\dot{M}_E^* \approx |\dot{M}_E|$. This much simplified calculation therefore shows that the E-ring complex could be partially maintained by accretion of the A-ring water grains as well as acquisition of the corresponding angular momentum. Thus, the E ring could have a lifetime much larger than 10^4 yr as estimated before.

In general, the plasmatisation model of the rings requires the presence of many water clusters in the ring system. To maintain the loss rate of $\dot{M} = 4 \times 10^6 \text{ g s}^{-1}$ and with photoionization as the only net ionization effect (timescale assumed to be $\sim 10^9 \text{ s}$), the average number density of the water clusters (assumed to contain 1,000 H_2O molecules each) would be $n_{\text{cluster}} \approx 10^7\text{--}10^8 \text{ cm}^{-3}$ in the ring plane with a thickness $\leq 10^2 \text{ km}$. Their existence can be tested by remote sensing methods or by *in situ* measurements. Neglecting the interception effect by the E ring, from equation (2) we have the total escape flux $\dot{N} \approx 10^5 \dot{M}_E / m_{\text{cluster}} \approx 2 \times 10^{24} \text{ s}^{-1}$. It is of interest to search for these water clusters and ions (after their re-ionization) in the magnetosphere of Saturn using appropriate ion and neutral detectors/analysers.

Let us consider the physical constraints on the generation of cluster ions. Hypervelocity impact experiments have shown that tiny condensates form as part of the impact ejecta (E. Grün, personal communication). One reason for their formation may be related to the supersaturation (and hence recondensation) effect in the adiabatically cooling outflow of the expanding impact vapour. Secondary cluster ions have also been detected and analysed in ion sputtering experiments²³. In the D-region of the Earth's ionosphere, water cluster ions are very abundant as a result of ion-neutral reactions²⁴. As for the required number density of such large cluster ions, we shall adopt the assumption that there is a population of cool plasma stored in the so-called 'free-zone' just above the ring plane²⁵. Taking the vertical displacement of the centre of the azimuthally symmetrical magnetic field to be $\sim 0.04 R_s$ (ref. 9), the total volume of the plasma disk so formed can be estimated to be $V \approx 1\text{--}8 \times 10^{28} \text{ cm}^3$. Assuming that the ring plasma contains mainly electrons and singly ionized cluster ions, we have the total mass loss rate caused by electronic recombination given as $\dot{M}_c \approx \alpha_e m_c (n_c^+)^2 V$, with α_e as the electronic recombination coefficient and n_c^+ as the number density

of the cluster ions in the free-zone. In laboratory studies²⁶, α_e for water cluster ions $\text{H}_3\text{O}^+(\text{H}_2\text{O})_n$ with $n = 0\text{--}6$ has been determined to be rather large ($\sim 10^{-5} \text{ cm}^3 \text{ s}^{-1}$). Now, equating $\dot{M}_c$ to $10^5 \dot{M}_E$ we find $n_c^+ \approx 2\text{--}11 \text{ cm}^{-3}$, which is reasonably small. (Taking the larger value of $n_c^+ \approx 11 \text{ cm}^{-3}$, which may be more appropriate for a thin plasma disk, the corresponding electronic recombination time is of the order of 10^4 s , which is a fraction of the bouncing period of the cluster ions and the planetary rotation period.)

In the estimate of the accretion probability P_E , it was assumed that the water clusters pass through the E ring only once. In fact, in the present scenario, the water clusters emitted at radial distance less than $2.34 R_s$ will be in periodic elliptical orbits, and they could return to the E-ring region repeatedly until lost by ionization or impact. For instance, for $R_o = 2.20 R_s$ (and $R_k = 3.70 R_s$), there will be about 10 returns if the ionization time is about $2 \times 10^6 \text{ s}$. This effect will increase the P_E value by a factor of 10 and leads to a reduction of the total recombination rate by a similar factor. This means the momentum coupling between the A ring and the E ring should require only a moderate amount of mass ejection from the A ring. Note that, besides water clusters, atoms and molecules like H, O and H_2O (and thus their ionized counterparts) could in principle perform the same function in mass and momentum transfer. The role of large water clusters and cluster ions is emphasized here, as they would imply a very small value of electron number density ($n_e \leq 10 \text{ cm}^{-3}$) along the magnetic flux tubes in the siphon flow process, whereas for atomic and molecular ions, $n_e \approx 10^3\text{--}10^4 \text{ cm}^{-3}$ is required. More quantitative work should clarify this point further.

Finally, we note that the proposed scenario of the plasmatisation and recondensation of the saturnian ring system is reminiscent of the model of Alfvén²⁷ which relates the mass distribution of the rings to the condensation process in a rotating plasma. The important difference is that the present model applies to the redistribution of ring matter after the initial formation of the ring system. We also note that other electromagnetic angular momentum transfer processes such as that discussed in the case of charged dust particles²⁸ could be important and complementary to the one described here.

Received 23 July 1985; accepted 14 January 1986.

1. Baum, W. A. *et al. Icarus* **47**, 84–96 (1980).
2. Larson, S. M., Fountain, J. W., Smith, B. A. & Reitsema, H. J. *Icarus* **47**, 288–290 (1981).
3. Smith, B. A. *et al. Science* **215**, 504–537 (1982).
4. Northrop, T. G. & Hill, J. R. *J. geophys. Res.* **87**, 6045–6051 (1982).
5. Northrop, T. G. & Hill, J. R. *J. geophys. Res.* **88**, 6102–6108 (1983).
6. Ip, W.-H. *J. geophys. Res.* **88**, 819–822 (1983).
7. Ip, W.-H. *J. geophys. Res.* **89**, 395–398 (1984).
8. Ip, W.-H. *Adv. Space Res.* **4**, 53–62 (1984).
9. Connerney, J. E. P., Ness, N. F. & Acuna, M. H. *Nature* **298**, 44–46 (1982).
10. Null, G. W., Lau, E. L., Biller, E. D. & Anderson, J. D. *Astr. J.* **86**, 456–468 (1981).
11. Lissauer, J. J., Peale, S. J. & Cuzzi, J. N. *Icarus* **58**, 159–168 (1984).
12. Henon, M. in *Planetary Rings* (ed. Brahic, A. 361–384 (Capadues, Toulouse, 1984).
13. Morfill, G. E., Fechtig, H., Grün, E. & Goertz, C. K. *Icarus* **55**, 439–447 (1983).
14. Ip, W.-H. *Icarus* **60**, 547–552 (1984).
15. Shimizu, H. *Moon Planets* **22**, 521–522 (1980).
16. Connerney, J. E. P. & Waite, J. H. *Nature* **312**, 136–138 (1984).
17. Ip, W.-H. in *Proc. Charged Particle Transport in Plasmas*, Spring College on Plasma Physics, Int. Center for Theoretical Physics, Trieste (World Scientific, Singapore, in the press).
18. Lissauer, J. J. & Cuzzi, J. N. *Astr. J.* **87**, 1051–1058 (1982).
19. Pang, K. D., Voge, C. C. & Rhoads, J. W. in *Planetary Rings* (ed. Brahic, A.) 607–614 (Capadues, Toulouse, 1984).
20. Morfill, G. E., Grün, E. & Johnson, T. V. *J. geophys. Res.* **88**, 5573–5579 (1983).
21. Lanzerotti, L. J. *et al. J. geophys. Res.* **88**, 8765–8770 (1983).
22. Haff, P. K., Eviator, A. & Siscoe, G. L. *Icarus* **56**, 426–438 (1983).
23. Lancaster, G. M., Honda, F., Fukuda, Y. & Rabalais J. *Am. chem. Soc.* **101**, 1951–1958 (1979).
24. Arnold, F., Krankowsky, D. & Marien, K. H. *Nature* **267**, 30–32 (1977).
25. Ip, W.-H. *Nature* **302**, 599–601 (1983).
26. Smith, D., Adams, N. G. & Church, M. J. *Planet. Space Sci.* **24**, 697–703 (1976).
27. Alfvén, H. *On the Origin of the Solar System* (Clarendon, Oxford, 1954).
28. Goertz, C. K. *et al. Nature* **320**, 141–143 (1986).

Random fluctuations versus Poynting–Robertson drag on interplanetary dust grains

Max K. Wallis

Department of Applied Mathematics and Astronomy, University College, Cathays Park, Cardiff CF1 1XL, UK

A particle moving through a sea of stochastic impulses gains energy from the sea; this is a long-established property of brownian motion, given theoretical backing in the fluctuation–dissipation theorem¹. However, this energy gain is generally neglected in describing circumstellar dust grains as slowly spiralling into the central star under Poynting–Robertson (P–R) drag. Stochastic forces on circumsolar grains arise from low-frequency fluctuations in surface charge and interplanetary fields, as has long been recognized^{2,3}, but previous studies have treated the mean orbital energy as constant^{4,5} or slowly decreasing⁶. I show here that stochastic heating can balance and indeed overwhelm the dissipative P–R drag for small grains ($\leq 1 \mu\text{m}$ scale radius). Furthermore, the consequent drift to the outer Solar System may cancel the strong collisional source inferred⁷ for micrometre and sub-micrometre fragments, explaining how their present distribution is stable in time. More generally, the stochastic heating mechanism combined with collisional fragmentation sets limits on dust accretion onto cool stars.

The stochastic forces $\mathbf{F}^-$ are formally included in the Langevin equation of motion of a mass m in the gravitational field of a large mass M_0 as

$$m\ddot{\mathbf{r}} = mM_0G\nabla\left(\frac{1}{r}\right) - \mathbf{F}^{\text{PR}} - \mathbf{F}^- \quad (1)$$

This generalized Kepler problem has been studied for other damping and stochastic forces; stationary solutions do exist in which, on average, energy gain compensates orbital damping, as demonstrated by Marshall and Claverie⁸ in the stochastic electrodynamic model of atoms. Following these authors, we start from the corresponding Fokker–Planck equation in the hamiltonian variables⁹ for the distribution function f

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial w_i} \left(-a_i f + \frac{1}{2} b_{ij} \frac{\partial f}{\partial w_j} \right) \quad (2)$$

This depends on assuming the fluctuations to be markovian and adopting the narrow-line approximation⁹. Then diffusive terms remain in the action variables, but not in the hamiltonian angles. The generalized force and diffusion coefficients are

$$a_j = \langle F_i^{\text{PR}} \partial w_j / \partial p_i \rangle, \quad b_{ij} = \tau^{-1} \langle \Delta w_i \Delta w_j \rangle \quad (3)$$

averaged around the slowly-changing Kepler orbits, b_{ij} being averaged also over the spectrum of fluctuations Δw arising from $\mathbf{F}^-$. In the Kepler problem, the w_j are equivalent to total energy $w_1 = p_i p_i / 2m - mM_0G/r$, total angular momentum $w_2 = \sqrt{(p_i p_i)}$ and angular momentum in a given direction $w_3 = p_z$; an alternative formulation^{4,6} employs the orbital elements instead of the above w_j .

Stationary solutions of equation (2) are not available for the Kepler problem⁸. However, the principle and the physical scale involved are illustrated by the one-dimensional analogue including source Q and loss terms L :

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial \varepsilon} \left(-af + \frac{1}{2} b \frac{\partial f}{\partial \varepsilon} \right) + Q(\varepsilon) - L(\varepsilon) \quad (4)$$

for purely circular orbits of scaled energy

$$\varepsilon = 2w_1/m = -2M_0G/r + p_\theta^2/m^2, \quad p_\theta^2/m^2 = M_0G/r \quad (5)$$

The Poynting–Robertson drag on a spherical grain of radius s (valid down to $0.3 \mu\text{m}$ radius in general, down to $0.1 \mu\text{m}$ for iron and graphite grains) is¹⁰

$$F^{\text{PR}} = -(s_0/s) M^0 G p_\theta / r^2 c, \quad \text{where } s_0 = 0.19 \mu\text{m}$$

so the damping force from equations (3) and (5) is

$$a = F^{\text{PR}} d\varepsilon/dp_\theta = -(s_0/s) 2p_\theta^2/m^2 r^2 c \quad (6)$$

The Fokker–Planck coefficient arising from the fluctuating electromagnetic impulse is

$$b = \tau^{-1} \langle \Delta \varepsilon \Delta \varepsilon \rangle = 4 \left\{ \frac{q\Omega}{mc} \right\}^2 P_{nn} p_\theta^2 / m^2 r^2 \quad (7)$$

where q is the grain charge, Ω is a constant (solar angular speed), c is the velocity of light and $P_{nn}(\omega; i, \theta)$ is the power spectrum of the fluctuating fields, to be averaged around orbital positions θ . Compare equation (7) with equation (20) of ref. 6, scaled by eccentricity; also $\langle \Delta a^2 \rangle$ of ref. 4 with arithmetical correction of a factor of 10 in equation (9). In the interplanetary medium, the mean P_{nn} is probably larger around orbits of small inclination i to the ecliptic plane than on highly inclined orbits (because reversals in the magnetic field direction which are confined to within $\pm 15^\circ$ contribute strongly). For the present spherically symmetric analogue, we take as representative the rough value⁴

$$P_{nn} = 5 \times 10^{48} \text{ G}^2 \text{ cm}^4 \text{ s} \quad (8)$$

corresponding to $1 \times 10^6 \text{ nT}^2 \text{ Hz}^{-1}$ power in field fluctuations ($1 \text{ nT} = 10^{-5} \text{ G}$). This value lies between values measured at small and large heliolatitudes, and is good to a factor of 2–3 (see ref. 6).

The stationary solution to equation (4) with no source or loss terms and vanishing at large r is trivial:

$$f \propto \exp \int (2a/b) d\varepsilon - 1 = e^{-\alpha \varepsilon} - 1 \quad \text{for } \alpha = -2a/b \quad (9)$$

since the parameters a and b have identical dependence ($\propto \varepsilon^3$). In terms of radial distance, $f \sim \exp(R/r) - 1$ with the scale radius

$$R = M_0 G \alpha = (s_0/s) (M_0 G m / q \Omega)^2 c / P_{nn} = R_0 \rho^2 s^3 \quad (10)$$

Taking the grain charge q to correspond to a 5-V potential ($q = (sV_0/300)$ e.s.u. for radius s in μm) and P_{nn} from equation (8), the scale R_0 is:

$$R_0 = 1.0 \text{ AU } \mu\text{m}^{-3} \text{ g}^{-2} \text{ cm}^{-6} \quad (11)$$

The exponential behaviour shows that stochastic heating is vigorous enough in $r > R$ to readily compensate the P–R damping; also, the inner region where damping dominates is smaller for smaller grains.

In the Solar System, grains are destroyed by evaporation near the Sun ($\varepsilon \approx -\infty$) or lost from loosely-bound orbits ($\varepsilon \approx 0$); comets and asteroids are sources of new grains, while both losses and gains result from fragmenting collisions¹¹. For simplicity, we follow ref. 7 by taking a source $Q = \Phi k |\varepsilon|^{k-1}$ with power index $k > 2$ and allow some loss $L_0 \delta(\varepsilon)$ at $\varepsilon = 0$. Then the stationary form of equation (4) integrates once to give the flux

$$J = af - \frac{1}{2} b \frac{\partial f}{\partial \varepsilon} = L_0 - \Phi |\varepsilon|^k \quad (12)$$

The distribution satisfying equation (12) is

$$f(\varepsilon) = 2 \int_\varepsilon^0 e^{\alpha(\varepsilon' - \varepsilon)} \frac{d\varepsilon'}{b(\varepsilon')} (L_0 - \Phi |\varepsilon'|^k) \quad (13)$$

where $b \sim |\varepsilon|^3$ from equations (5) and (7). Formally, the loss integral is singular at $\varepsilon = 0$ and we should take $L_0 \rightarrow 0$. This is in part a consequence of the diffusion approximation, overcome in stochastic models having finite step size (as for comet orbit

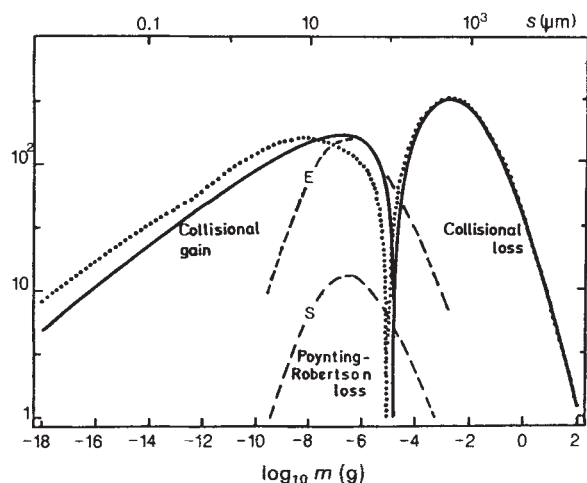


Fig. 1 Net gain and loss rates (arbitrary units) due to collisions evaluated in ref. 7 at 1 AU for two different grain density models (solid line from interplanetary dust collection; dotted line from lunar microcrater counts), compared with the differential P-R loss⁷ for spherical grains (S) and estimated enhanced loss for elongated grains (E), which assume spatial density variation as $r^{-1.3}$. The upper size scale is for spherical grains of density 2.4 g cm^{-3} .

'diffusion'; see, for example, ref. 12), but in part results because $b \sim r^{-2}$, so that particles accumulate in the outer Solar System until removed by accretion or fragmenting collisions. Indeed, at a distance of order 100 AU, where the solar wind degrades or terminates, b would fall off still faster and produce stronger accumulation. As this aspect is not modelled here, dropping the loss integral and taking f_0 as finite is legitimate. Thus we have

$$f(\varepsilon) = e^{\alpha\varepsilon} \left\{ f_0 - \frac{2}{\beta} \Phi \int_0^{|\varepsilon|} s^{\alpha y} y^k \frac{dy}{y^3} \right\} \quad (14)$$

where the constant $\beta = b|\varepsilon|^{-3}$ depends strongly on grain size ($\beta \propto s^{-4}$). No cutoff in the source is needed at small ε for $k > 2$. Evidently for $k = 3$, the integral takes the simple form of equation (9). An inner boundary condition could be imposed corresponding to evaporation of grains near the Sun, but all we need is to ensure that $J < 0$ at 0.1 AU. The explicit solution available for $k = 3$ is

$$f = f_0 + \frac{2}{\alpha\beta} \Phi \exp\left(\frac{6 \text{ AU}}{r} s [\mu\text{m}]^3\right)^{-1} \quad (15)$$

after using equations (10) and (11), and setting ρ equal to 2.4 g cm^{-3} . The combined multiplicative factors do not vary strongly with grain size: the dependence is roughly $\Phi \sim s^{-2}$ for $s < 2 \mu\text{m}$ (or 10^{-10} g ; see Fig. 1), $\Phi/\alpha\beta \sim s^{-1}$. But the exponential factor in equation (15) can be huge, for example $\exp(6,000)$ for $10\text{-}\mu\text{m}$ grains, because diffusion is very weak and sources and sinks dominate. Thus, the zodiacal light indicates that $10\text{-}100\text{-}\mu\text{m}$ grains are confined to the ecliptic plane, as collisions and P-R forces dominate⁷. For $0.2\text{-}\mu\text{m}$ grains, on the other hand, the diffusive solution varies relatively slowly, as $\exp(0.6 \text{ AU}/r)$, and diffusive processes dominate. The boundary of applicability of the simple stochastic solution (14) is thus plausibly $s \approx 0.5 \mu\text{m}$ (10^{-12} g).

The scales are changed if damping forces operate that exceed the classical P-R force. Sputtering of grains by the solar wind, the direction of which is a few degrees from radial, gives¹³ a pseudo-P-R force with doubled parameters a (of equation (6)) and R_0 (of equation (11)) for prograde orbiting dust particles. A second variant covers highly-elongated grains partially aligned by the magnetic field, which scatter photons anisotropically and

suffer a mean tangential force many times higher than the P-R force¹⁴. In the absence of particular analysis and knowledge of actual shapes, this may justify increasing P-R loss rates by an order of magnitude to balance the net collisional gains, as in Fig. 1 (dashed line E). Because the enhancement depends little on grain size¹⁴, it cannot explain the increasing divergence evident in Fig. 1 between gains and losses of grains smaller than a few micrometres.

We have demonstrated that the stochastic energy addition ignored in previous studies can, in principle, overcome the spiralling in of dust grains due to P-R drag. The one-dimensional model explored here should be an adequate description of the problem. Variation with orbit eccentricity and inclination are not directly important, because orbital periods are unaffected and mean stochastic rates depend only weakly on them⁶. Insofar as P_{nn} is up to 3 times higher for low-orbit inclinations, diffusive speeds may be higher and densities lower by a factor $\sqrt{3}$ in the ecliptic plane. The systematic electromagnetic effects that strongly deflect grains smaller than $\sim 0.3 \mu\text{m}$ in the inner Solar System reverse sign with solar polarity every 11 yr; diffusive behaviour is thus valid outside Jupiter's orbit, although periodic effects would be evident in the Earth's neighbourhood. The μm -sized grains released mainly in the ecliptic plane would diffuse initially in inclination⁴⁻⁶. Grains of a few micrometres have mean P-R drag reduced by energy gains, and so continue to orbit along P-R spirals; their space density depends primarily on diffusion out of the plane, not modelled here. Observations of the particle spectrum at 1 AU do show evidence of a new mechanism operating on submicrometre grains; the break or 'knee' at around $0.5\text{-}1 \mu\text{m}$ in the size distribution¹¹ would correspond to the change from source-dominated to stochastic diffusive-type solutions. The steeper spectrum of grains of $s < 1 \mu\text{m}$ inferred by Lamy and Perrin¹⁵, if confirmed to exist, would constitute the diffusion-dominated population.

One longstanding issue concerning the zodiacal cloud is the nature of the source which replaces the estimated $10\text{-}20$ tonnes s^{-1} lost within 1 AU through fragmenting collisions¹⁶. Short-period comets have been thought inadequate, but there are uncertainties in the estimates and other possible sources exist for the required sub-millimetre meteoroids⁷. The more pressing issue concerns removal of the fragments from collisions that we know must occur. Whipple¹⁶ supposed that one third of the fragments would go immediately into hyperbolic orbits, but found that the removal rate through P-R forces is only 5–10% of the production rate. Grün *et al.*⁷ estimate a 50% loss to hyperbolic orbits but, excluding the counteracting stochastic effects, still only 5% through P-R drag (see Fig. 1). If, however, P-R forces were ~ 10 times stronger, as appears plausible from the calculations for cylindrical grains¹⁴, they could remove the excess in the $10\text{-}100\text{-}\mu\text{m}$ range ($10^{-8}\text{-}10^{-5} \text{ g}$). Nevertheless, the strong excess production of micrometre and smaller grains remains (Fig. 1), so that their density would be increasing over the relatively short timescale of $10,000 \text{ yr}$. The present analysis identifies a new removal mechanism that can solve this conundrum. The stochastic forces add energy on the average, ejecting grains towards the outer Solar System. These grains might accumulate there over long periods of time and eventually perhaps accrete onto meteoroids and comets.

I thank Trevor Marshall and Ajay DasGupta for helpful discussions.

Received 21 June 1985; accepted 13 January 1986.

1. Kubo, R. *Rep. Prog. Phys.* **29**, 255–285 (1966).
2. Parker, E. N. *Astrophys. J.* **139**, 951–958 (1963).
3. Morfill, G. E. & Grün, E. *Planet. Space Sci.* **27**, 1283–1292 (1979).
4. Consolmagno, G. *Icarus* **38**, 398–410 (1979).
5. Barge, P., Pellat, R. & Millet, J. *Astr. Astrophys.* **115**, 8–19 (1982).
6. Wallis, M. K. & Hassan, M. H. A. *Astr. Astrophys.* **151**, 435–441 (1985).
7. Grün, E., Zook, H. A., Fechtig, H. & Giese, R. H. *Icarus* **62**, 244–272 (1985).
8. Marshall, T. W. Claverie, P. *J. math. Phys.* **21**, 1819–1825 (1980).
9. Marshall, T. W. *Physica* **103A**, 172–182 (1980).

10. Burns, J. A., Lamy, Ph. L. & Soter, S. *Icarus* **40**, 1–48 (1979).
11. Leinert, C., Röser, S. & Buitrago, J. *Astr. Astrophys.* **118**, 345–357 (1983).
12. Yabushita, S. *Astr. Astrophys.* **85**, 77–79 (1980).
13. Mukai, T. & Yamamoto, T. *Astr. Astrophys.* **107**, 97–100 (1982).
14. Voschinnikov, N. V. & Il'in, V. B. *Pis'ma Astr. Zh.* **9**, 188–192 (1983). (*Soviet Astr. Lett.* **9**, 101–103; 1983).
15. Lamy, Ph. L. & Perrin, J. M. in *Solid Particles in the Solar System* (eds Halliday, I. & McIntosh, B. I.) 75–80 (Reidel, Dordrecht, 1980).
16. Whipple, F. L. in *Zodiacal Light and the Interplanetary Medium* 409–326 (NASA SP-150 1967).

Geomagnetic reversal spurts and episodes of extraterrestrial catastrophism

Poorna C. Pal* & Kenneth M. Creer†

* Department of Geology & Mineral Sciences, University of Ilorin, PMB 1515, Ilorin, Nigeria

† Department of Geophysics, University of Edinburgh, Edinburgh EH9 3JZ, UK

The possibility that the observed 26–32-Myr periodicity in the records of extinction events¹ and astrobleme ages² was caused by periodic episodes of physical bombardment of the planetary surface has raised the likelihood that such an astronomical clock also affects geomagnetic reversals^{3,4}. However, the detection⁵ of a comparable 30-Myr harmonic in the geomagnetic reversal record has now been found⁶ to be an accident of data length while the earlier-reported⁷ 32–34-Myr harmonic of the reversal record may well be an artefact of the data base. Here we examine the temporal relation between geomagnetic reversals and those catastrophic episodes which occurred during periods of frequent reversals. As geomagnetic history comprises long intervals of frequent reversals, alternating with intervals of fixed polarity, our rationale is that the geomagnetic field should be most susceptible to catastrophism during its relatively unstable periods of frequent reversals. We find spurts in the frequency of geomagnetic reversals at approximately 30-Myr intervals, correlated with catastrophic episodes. These spurts can be seen as the geomagnetic signatures of terrestrial catastrophism, and we tentatively ascribe them to enhanced core turbulence during episodes of bombardment. The absence of reversal spurts during long intervals of fixed polarity is attributed to the failure during these intervals of the geomagnetic dynamo's bi-stable oscillations.

Polarity reversals of the geomagnetic field have occurred often in the geological record⁸, although with varying frequency. The durations of the individual normal (N) and reversed (R) regimes have thus ranged from 0.02–0.03 Myr for the early Pleistocene (Olduvai) N-polarity intervals ('subchrons') to 20–70 Myr for the late Palaeozoic R- and early and late Mesozoic N-polarity intervals (Kiaman, Balkan and Gubbio 'superchrons'). Figure

1 shows the variation of reversal frequency since the late Palaeozoic as seen through 5-Myr rectangular windows at 2.5-Myr intervals, based on a largely marine-derived magnetic record for the post-Callovian period⁸ and magnetostratigraphic data⁹ for earlier times. A rhythmic variation in reversal frequency with time is readily apparent here, with periods of few or no reversals (the fixed-polarity superchrons) lasting 20–70 Myr, alternating with equally long periods of frequent reversals (the mixed-polarity superchrons). The N and R interval lengths tend to increase towards the fixed polarity superchrons; however, the likely significance of this trend, reported in several recent studies^{10–12}, is obscured by an abrupt rise in reversal frequency following the late Triassic–early Jurassic N-polarity superchron. Similarly, although the lengths of the superchrons appear to decrease with age, this observation may be dramatically affected by any improvement in resolution and detection threshold, particularly as the Norian–Callovian reversal sequence remains far from complete.

The time distribution of reversals is thus enigmatic. Fourier spectrum¹³, maximum entropy^{9,14}, Walsh sequence⁷, minimum dispersion⁵ and autocorrelation¹¹ analyses, for example, have been used to demonstrate the presence in the reversal spectrum of harmonics with periods of approximately 700, 250–300, 100–120, 45–70, 30–34 and 15 Myr. The incompleteness of the record casts doubt on the significance of the longer-period harmonics cited; even the proposed 30- and 15-Myr harmonics are controversial^{6,15,16}. As for the 32–34-Myr harmonic⁷, the tentativeness of the database and the inclusion of fixed-polarity superchrons in the analysis (one of which is of ~35-Myr duration) make its identification rather suspect. Its correlation with the 26–32-Myr periodicity of catastrophic episodes is hindered by the failure to identify the associated geomagnetic signatures, as these episodes occurred during mixed- as well as fixed-polarity superchrons. From a theoretical standpoint, the postulated^{12,17} random nature of the geomagnetic reversal process rules out the presence of well-defined harmonics in the reversal frequency spectrum. Thus, if there is a link between geomagnetic reversals and global catastrophe^{5,7}, it is not revealed as a common periodicity.

What then is the geomagnetic signature of a catastrophic episode? One possibility is that the reversal record is totally free of any such signatures. Another is that catastrophic episodes are expressed only in mixed-polarity superchrons. The post-Callovian record⁸, the most reliable segment of the data in Fig. 1, comprises two such superchrons: the post-Santonian sequence (from 83 Myr to the present) and the Oxfordian–Barremian sequence (~163–119 Myr). The first of these is analysed here, the latter having been too brief for a detailed analysis.

Figure 2 shows the variations in reversal frequency since early Campanian based on the Lowrie–Alvarez (LA-81) timescale¹⁸ as seen through a 15-Myr window sliding at 1-Myr intervals. The reversal rate curve (a) shows clear fluctuations, with peaks in reversal frequency at about 65–75, 35–45 and <8–12 Myr

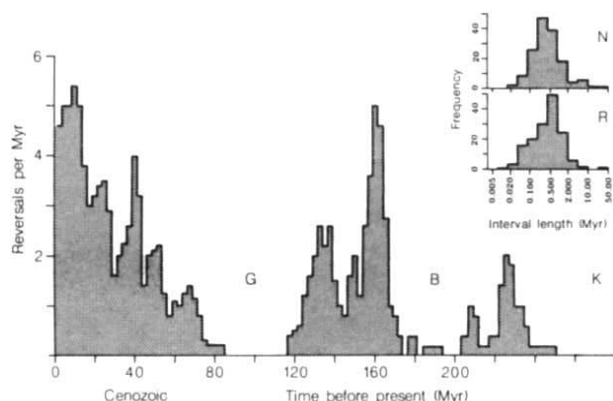


Fig. 1 Geomagnetic reversal frequency as a function of geological time since the late Permian computed for successive 5-Myr segments in 2.5-Myr overlaps using the timescale of Harland *et al.*⁸ for the post-Callovian (≤ 163 Myr) interval and magnetostratigraphic data⁹ for earlier times. The inset shows that the N and R interval lengths in this chronology are log-normally distributed (about means of 0.387 and 0.367 Myr, respectively), implying that intervals shorter than about 0.1 Myr have been almost as uncommon as those longer than about 2 Myr. Long periods of few or no reversals (fixed-polarity superchrons) are indicated by letters: G, Gubbio (Mercatton) N interval; B, Balkan (Graham) N interval; K, Kiaman R interval.

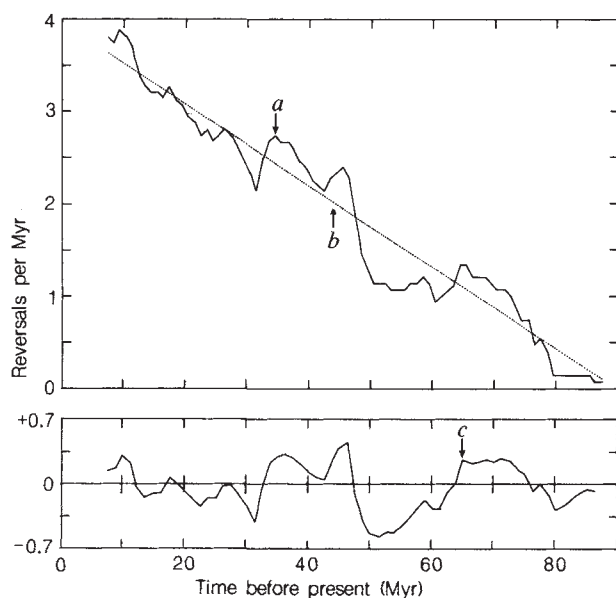


Fig. 2 Geomagnetic reversal frequency (*a*) during the present (post-Santonian) mixed-polarity superchron, obtained by viewing the LA-81 timescale¹⁸ through a 15-Myr rectangular window, chosen in order to remove the 15-Myr harmonic of Mazaud *et al.*¹¹. The overall trend of decreasing frequency with age is modelled here using a straight line fit (*b*): $RF(T) = 3.96 - 0.0445 T$, where $RF(T)$ is reversal frequency (per Myr) at any time T (Myr) during the past 90 Myr. The residual curve (*c*) is obtained by removing this linear fit from the reversal rate curve and clearly emphasizes the rhythmicity in the reversal rate curve, with spurts during the intervals 65–75, 35–45 and 8–12 Myr.

superposed on a linear decrease with age. When we detrend this by removing a straight line fit (*b*), the residual curve (*c*) is found from the Fourier and maximum entropy analyses (Fig. 3) to be dominated by a 26–32-Myr harmonic. While the data length analysed here is too short to be able to call this a basic harmonic of the reversal record, these results do confirm the presence of three reversal spurts separated by ~30 Myr.

The assumption that these 30-Myr spurts in the reversal frequency are otherwise masked by a 15-Myr harmonic, which prompted our selection of a 15-Myr window, is amply demonstrated by the following analysis of the post-Santonian reversal record. Figure 4 shows the reversal frequency histogram obtained when the LA-81 scale¹⁸ is viewed through a 5-Myr window, along with the corresponding power spectrum (inset). An overwhelming proportion of the power ($\geq 80\%$) is accounted for by the 45-Myr and longer-period harmonics; the remainder is dominated by the 28–36- and 14–18-Myr harmonics. Also shown in Fig. 4 (dotted curve) is the reversal rate curve obtained by deleting harmonics of period ≤ 18 Myr. Reversal spurts centred at about 12, 38 and 67 Myr appear distinctly in this curve, corroborating the reversal rate curve obtained in Fig. 2 by a 15-Myr smoothing.

Our results thus suggest that the presently continuing superchron of frequent reversals is characterized by three clearly defined reversal spurts spaced about 30 Myr apart. If the geomagnetic reversal process were characterized by a 30-Myr periodicity, we would expect to see a reversal spurt ~30 Myr before the oldest of these three; that is, in Cenomanian or late Albian times. There was no such spurt; indeed, the field remained fixed in the N-polarity state throughout the 35-Myr Aptian–Santonian interval⁸. The preceding Oxfordian–Barremian mixed-polarity superchron did contain two such spurts at 130–140 and 155–165 Myr (ref. 5) (see Fig. 1), but this superchron was too short to allow detailed analysis.

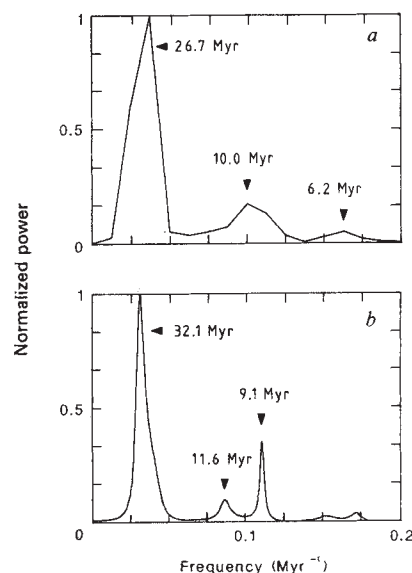


Fig. 3 Fourier (*a*) and maximum entropy (*b*) spectra for the residual curve of Fig. 2. Note the dominance of a 26.7–32.1-Myr harmonic, reflecting the removal of shorter-period harmonics by the 15-Myr smoothing, and of longer-period harmonics by linear detrending. The 6.2–11.6-Myr periods in these spectra indicate leftovers from the smoothing process used to compute the reversal-rate curve shown in Fig. 2.

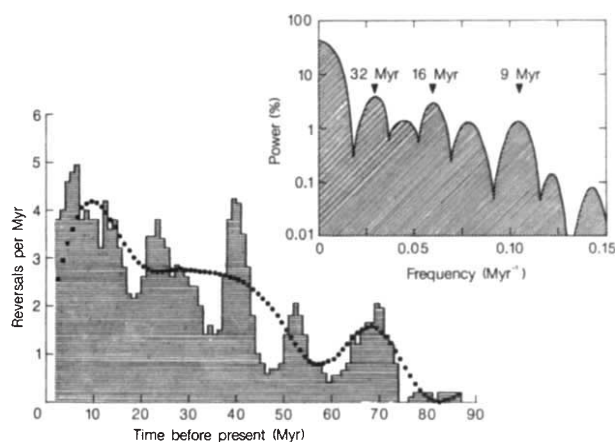


Fig. 4 The reversal frequency plot obtained by viewing the LA-81 scale¹⁸ through a 5-Myr rectangular window, before (histogram) and after (dotted curve) filtering the 18-Myr and shorter-period harmonics. The inset shows the power spectrum obtained from the unfiltered data.

Our finding that reversal spurts indeed characterized the mixed-polarity superchrons, occurring approximately once every 30 Myr since ~165 Myr except during the Cretaceous N-polarity superchron, prompts a search for a chronological link between these spurts and suspected catastrophic episodes. Evidence for the occurrence of global-scale catastrophic events includes impact craters, tektites and impactites signifying bombardment episodes, coeval geochemical anomalies implying an extraterrestrial origin for the bombardment, and the chronologically correlated extinction events as their attendant environmental repercussions. The terminal-Cretaceous and

Table 1 Correlation of reversal spurts and catastrophic episodes

Geological time	Reversals		Evidences of catastrophism				
	Polarity superchron	Reversal spurt	Astroblemes	Impactites	Tektites	Geochemical anomalies	Extinction event
Late-middle Miocene	Mixed	×	×	×	×		×
Early Oligocene-late Eocene	Mixed	×	×	×	×	×	×
Palaeocene-late Cretaceous	Mixed	×		×	×	×	×
Middle Cretaceous	Fixed		×	×			×
Early Cretaceous	Mixed	×	×	×			×
Late Jurassic	Mixed	×	×				×
Early-middle Jurassic	Fixed		×				×
Late Triassic	Mixed	?	×				×
Late Permian	Fixed					×	×

terminal-Eocene catastrophic episodes are thus inferred from the presence of iridium and osmium anomalies associated with extinction events^{19,22}, and from the discoveries of coeval astrobleme structures², impactites²³ and microtektites^{24,25}. The middle-late Miocene boundary is similarly characterized by a major extinction event¹, several impact structures² and the East European tektite strewn field^{25,26}. Palaeontological and crater-age data also indicate catastrophic episodes in the middle (Cenomanian) and early (Hauterivian) Cretaceous, late (Callovian-Tithonian) and early-middle (Bajocian-Pliensbachian) Jurassic, late Triassic (Norian) and late Permian (Guadalupian-Dzhulfian), yielding an estimate of about 30 Myr for the periodicity of catastrophic episodes since the end of the Palaeozoic¹⁻³.

Table 1 summarizes a chronological correlation of reversal spurts and catastrophic episodes since the late Palaeozoic. Except for the late Permian and middle Cretaceous episodes, which occurred during the Permian R and Cretaceous N fixed-polarity superchrons respectively, all the other catastrophic episodes are associated with reversal spurts. Questions remain about the presumed periodicity in the palaeontological record²⁷, but Raup's²⁸ re-analysis has confirmed its statistical significance. Similarly, tektites are not universally regarded²⁹ as indicative of extraterrestrially caused catastrophism, although the inference that they represent the reworking of terrestrial material attendant upon bolide collisions is widely accepted. Note that Fig. 1 suggests an element of cyclicity in the occurrences of the alternate mixed- and fixed-polarity superchrons. The power spectrum in Fig. 4 reflects this, with a concentration of $\geq 80\%$ of the total power in the 45-Myr and longer-period harmonics. Superposed on this pattern, the reversal spurts are so located in the mixed-polarity superchrons as if a basic 30-Myr harmonic in the reversal spectrum, extending back to at least 165 Myr and probably to the late Permian, is masked during the fixed-polarity superchrons. In view of the correlations shown in Table 1, we suggest that the approximately periodic recurrences of catastrophic episodes caused reversal spurts during the mixed-polarity superchrons. It would thus seem likely that these reversal spurts had little to do with any stochastic behaviour intrinsic to the geomagnetic dynamo mechanism and possibly represent an additive factor superposed on the otherwise random occurrences of reversals.

The geomagnetic field is thought to arise from magnetohydrodynamic motions in the Earth's fluid outer core³¹, and reversals in polarity are likely to be caused by core turbulence³⁰. The correlation of reversal spurts and catastrophic episodes can be explained in terms of enhanced turbulence in the core motions resulting from intensified physical bombardments of the planetary surface during the catastrophic episodes. The observed³²⁻³⁴ diminution of field strength and departure from axial symmetry during a polarity reversal is consistent with a core turbulence model for reversals. Similarly, evidence from lunar magnetism³⁵ supports our suggestion that extraterrestrially

induced catastrophism can drastically affect geomagnetic behaviour.

If catastrophic episodes indeed caused reversal spurts, we need to explain why fixed-polarity superchrons remained unaffected by them and why the spurts are of such long duration (relative to the duration of the catastrophic episodes). The first of these results from the likelihood that the geomagnetic dynamo is locked into a frozen state during a fixed-polarity superchron, implying an occasional failure of its random bi-stable oscillations¹², from which a catastrophism-induced resonance cannot readily free it. As for the second, the explanation lies partly in the broadening effect of smoothing (Fig. 2) and filtering (Fig. 4), and partly in the possibility that the corresponding catastrophic episodes comprised multiple pulses, not isolated events. The latter suggestion is supported by the discovery²⁴ of several microtektite layers spanning a 10-Myr interval near the Eocene-Oligocene boundary.

We thank David Raup and Timothy Lutz for constructive reviews and Bill Napier and Victor Clube for helpful discussions. P.C.P. also acknowledges the facilities provided by Universidad Nacional Autonoma de Mexico and University of California, Los Angeles.

Received 12 August 1985; accepted 23 January 1986.

1. Raup, D. M. & Sepkoski, J. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 801-805 (1984).
2. Alvarez, W. & Muller, R. A. *Nature* **308**, 718-720 (1984).
3. Rampino, M. R. & Strothers, R. B. *Nature* **308**, 709-712 (1984); *Science* **226**, 1427-1431 (1984).
4. Clube, S. V. M. & Napier, W. M. *Nature* **311**, 635-636 (1984); *Mon. Not. R. astr. Soc.* **208**, 575-588 (1984); *Earth planet. Sci. Lett.* **57**, 251-262 (1982).
5. Raup, D. M. *Nature* **314**, 341-343 (1985).
6. Lutz, T. M. *Nature* **317**, 404-407 (1985).
7. Negi, J. G. & Tiwari, R. K. *Geophys. Res. Lett.* **10**, 713-716 (1983).
8. Harland, W. B. *et al.* *A Geologic Time Scale* (Cambridge University Press, 1982).
9. Irving, E. & Pullaiah, G. *Earth Sci. Rev.* **12**, 35-64 (1976).
10. Lowrie, W. & Kent, D. V. *Earth planet. Sci. Lett.* **62**, 305-313 (1983).
11. Mazaud, A., Laj, C., de Seze, L. & Verosub, K. L. *Nature* **304**, 328-330 (1983).
12. McFadden, P. L. & Merrill, R. T. *J. geophys. Res.* **89**, 3354-3362 (1984).
13. Crain, I. K., Crain, P. L. & Plaunt, M. S. *Nature* **223**, 282-283 (1969).
14. Ulyrich, T. J. *Nature* **235**, 218-219 (1972).
15. McFadden, P. L. *Nature* **311**, (1984).
16. Mazaud, A., Laj, C., de Seze, L. & Verosub, K. L. *Nature* **311**, 396 (1984).
17. Cox, A. *Phys. Earth & planet. Int.* **24**, 178-190 (1981).
18. Lowrie, W. & Alvarez, W. *Geology* **9**, 393-397 (1981).
19. Alvarez, L. W. & Alvarez, W. *Science* **223**, 1174-1186 (1983).
20. Ganapathy, R. *Science* **216**, 885-886 (1982); *Science* **209**, 921-922 (1980).
21. Kyte, F. T., Smit, J. & Wasson, J. T. *Earth planet. Sci. Lett.* **73**, 183-195 (1985).
22. Luck, J. M. & Turekian, K. K. *Science* **222**, 613-615 (1983).
23. Bohor, B. F., Foord, E. E., Modreski, P. J. & Triplehorn, D. M. *Science* **224**, 867-868 (1984).
24. Keller, G., D'Hondt, S. & Vallier, T. L. *Science* **221**, 150-152 (1983).
25. Glass, B. P., Swincki, M. B. & Zwart, P. A. *Proc. 10th lunar planet. Sci. Conf.* Vol. 3, 2535-2545 (Pergamon, Oxford, 1979).
26. Barnes, V. F. in *Tektites* (ed. O'Keefe, J. A.) 25-50 (University of Chicago Press, 1963).
27. Hoffman, A. *Nature* **315**, 659-661 (1985).
28. Raup, D. M. *Nature* **317**, 384-385 (1985).
29. Crawford, A. R. *Q. Jl R. astr. Soc.* **26**, 53-55 (1985); *Geol. Mag.* **116**, 261-283 (1979).
30. Olson, P. *Phys. Earth planet. Inter.* **33**, 260-274 (1983).
31. Parker, E. N. *Cosmical Magnetic Fields* (Clarendon, Oxford, 1979).
32. Prevot, M., Maniken, E. A., Gromme, C. S. & Coe, R. S. *Nature* **316**, 230-234 (1985).
33. Dagley, P. & Lowley, E. *Geophys. J. R. astr. Soc.* **36**, 577-598 (1974).
34. Pal, P. C. *J. Ind. Acad. Geosci.* **16**, 81-96 (1973).
35. Runcorn, S. K. *Q. Jl R. astr. Soc.* **26**, 137-146 (1985); *Nature* **304**, 589-596 (1983).

Aluminium/silicon disordering and melting in sillimanite at high pressures

T. J. B. Holland & M. A. Carpenter

Department of Earth Sciences, University of Cambridge,
Downing Street, Cambridge CB2 3EQ, UK

Naturally occurring polymorphs of Al_2SiO_5 (andalusite, kyanite and sillimanite) have assumed a special significance for geologists because of their value as indicators of the pressures (P) and temperatures (T) experienced by metamorphic rocks^{1,2}, and for materials scientists because the binary system Al_2SiO_5 – Al_2O_3 contains the important refractory mineral, mullite³. The equilibrium stability field of sillimanite in natural rocks and the process by which it can be transformed to mullite for ceramic materials (mullitization) have received particular attention because of the role of cation ordering ($\text{Al}^{3+}/\text{Si}^{4+}$) in stabilizing these structures^{1–7}. Here we present the results of annealing experiments undertaken to determine the disordering behaviour of sillimanite crystals at high confining pressures. At $P \approx 18$ – 20 kbar and $T \approx 1,300$ – $1,700$ °C, they undergo partial melting to give a SiO_2 -rich melt. The residual crystals become progressively more disordered as they become enriched in Al_2O_3 until they transform to a higher-symmetry structure. Our observations shed new light on previous thermochemical studies of Al/Si order-disorder in sillimanite and on the phase relationships between sillimanite and mullite.

In principle, sillimanite should show an Al/Si disordering transformation at high temperature^{8,9}. To study such a transformation experimentally, however, high pressures must be applied to suppress the breakdown reaction sillimanite $\rightarrow$ mullite + SiO_2 and incongruent melting^{10,11}. The disordering reaction would result in a halving of the c unit-cell repeat and a space group change from Pbnm to Pbam ^{6,10,12}, and should be readily detect-

able by single crystal diffraction methods. Reflections of the type $l = \text{odd}$ are absent in diffraction patterns from disordered crystals^{6,10,12}, but are present and sharp in ordered sillimanite diffraction patterns². (The incommensurate superlattice of mullite gives pairs of reflections about these $l = \text{odd}$ positions^{2,4}). The original intention of the present experiments was to bracket the $\text{Pbam} \rightleftharpoons \text{Pbnm}$ transformation in sillimanite by the use of high- P and $-T$ annealing experiments and transmission electron microscopy (TEM).

Natural sillimanite crystals from Connecticut, USA (Cambridge mineral collection no. 3076) were annealed in sealed Pt capsules at pressures of 18–20 kbar and temperatures of 1,300–1,690 °C for 0.3–10 h in a piston-cylinder apparatus. The product crystals were examined at 100 kV in an AEI EM6G or a JEM100C electron microscope, as crushed grains deposited on carbon film and as ion-beam-thinned slices. The starting material had no pervasive microstructure and gave strong, sharp $l = \text{odd}$ reflections in electron diffraction patterns. The product crystals contained abundant dislocations, both isolated and in subgrain boundaries (Fig. 1), features resembling antiphase boundaries (APBs) in dark-field images formed with $l = \text{odd}$ reflections (Fig. 1a, b) and small inclusions on a scale of ~ 200 – $2,000$ Å (Fig. 1c–e). In addition, the intensity of $l = \text{odd}$ reflections in the electron diffraction patterns showed considerable variation. The observations are summarized in Table 1.

The APBs were commonly observed to end at dislocations (Fig. 1a), but in some cases formed quite regular domain textures (Fig. 1b). Antiphase domains (APDs) with a stacking vector of $\frac{1}{2}[001]$ would be expected to appear during an ordering transformation from Pbam to Pbnm in sillimanite. The same stacking mistakes can also be generated by the passage of partial dislocations, however, since deformed sillimanites commonly contain $[001]$ -type dislocations which dissociate into partials $\frac{1}{2}[001] + \frac{1}{2}[001]$ (refs 13–16). Stacking faults between partials of this type also have a displacement vector of $\frac{1}{2}[001]$ and cause the juxtaposition of Si next to Si and Al next to Al where, in defect-free regions, the Al and Si would alternate in a fully

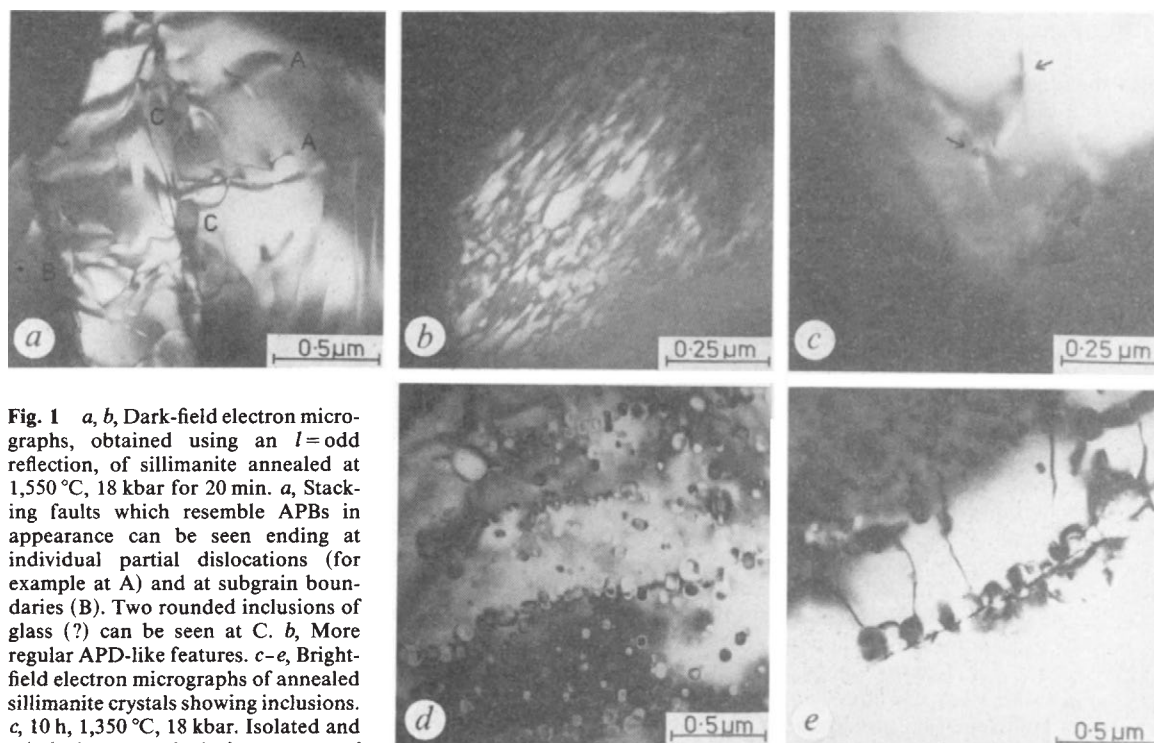


Fig. 1 a, b, Dark-field electron micrographs, obtained using an $l = \text{odd}$ reflection, of sillimanite annealed at 1,550 °C, 18 kbar for 20 min. a, Stacking faults which resemble APBs in appearance can be seen ending at individual partial dislocations (for example at A) and at subgrain boundaries (B). Two rounded inclusions of glass (?) can be seen at C. b, More regular APD-like features. c–e, Bright-field electron micrographs of annealed sillimanite crystals showing inclusions. c, 10 h, 1,350 °C, 18 kbar. Isolated and relatively scarce inclusions, some of which are associated with dislocations (arrowed). d, 30 min, 1,680–1,690 °C, 18 kbar. Abundant inclusions, interpreted as consisting of silica-rich glass. e, 30 min, 1,680–1,690 °C, 18 kbar. Large inclusions apparently formed by nucleation at a subgrain boundary. Electron diffraction patterns from the inclusion-free zone have $l = \text{odd}$ reflections absent.

Table 1 Annealing conditions and microstructures of heat-treated sillimanite crystals

Run no.	T (°C)	P (kbar)	Annealing time (h)	Product*
S10	1,680–1,690	20	0.5	~15–20% inclusions; <i>l</i> = odd reflections generally weak, but absent in inclusion-free zones.
S5	1,650	18	0.5	Abundant inclusions; <i>l</i> = odd reflections generally weak, but absent in inclusion-free zones.
S7	1,620	20	0.66	Abundant APBs; many inclusions; <i>l</i> = odd strong.
S2	1,550	18	0.33	Many inclusions; abundant APBs; <i>l</i> = odd strong.
S1	1,450	20	1.0	Inclusions scarce; abundant APBs; <i>l</i> = odd strong.
S4	1,350	18	10	Relatively scarce inclusions; abundant APBs; <i>l</i> = odd strong.
S3	1,350	20	0.5	Relatively scarce inclusions; abundant APBs; <i>l</i> = odd strong.
S12	1,300	20	9	Inclusions scarce; abundant APBs; <i>l</i> = odd strong.

* All products also contained abundant dislocations.

ordered manner^{13–16}. Once formed, these stacking defects would be free to migrate, and configurations resembling true APD textures might result. The common association of APB-like features with dislocations suggests that a deformation origin is the most likely explanation for their occurrence in the present case.

We interpret the small inclusions in Fig. 1c–e as silica-rich glass resulting from partial melting of the sillimanite crystals. They are typically associated with dislocations and subgrain boundaries, suggesting heterogeneous nucleation of the melt. In some cases they are slightly elongate, with flat sides parallel to *c* and rounded ends. Their amorphous character has been deduced from the absence of any diffraction maxima other than those of sillimanite and from the lack of diffraction contrast within the inclusions. Small fragments of isotropic material were also observed by high-magnification optical microscopy in the 1,680–1,690 °C run product.

Crystals from runs at 1,300–1,550 °C gave diffraction patterns with strong *l* = odd reflections and had relatively few inclusions, which were less homogeneously distributed (Fig. 1c). Crystals annealed at 1,620 °C had a higher density of inclusions but the *l* = odd reflections were still strong. Crystals from the two highest-temperature runs contained significantly more inclusions (up to 15–20% at 1,680–1,690 °C; Fig. 1d), and diffraction patterns from these contained obviously weakened (but still sharp) *l* = odd reflections. One relatively inclusion-free region from the 1,680–1,690 °C run (Fig. 1e) gave diffraction patterns with *l* = odd absent. In this case, a subgrain boundary had evidently provided suitable sites for heterogeneous melt nucleation; diffusion of Si and Al between the liquid and residual crystal then occurred on a sufficient scale to leave an inclusion-free zone. Small inclusions of glass were spread more homogeneously throughout the rest of the crystal. It is clear, therefore, that partial melting at the highest temperatures was accompanied by substantial Al/Si disordering.

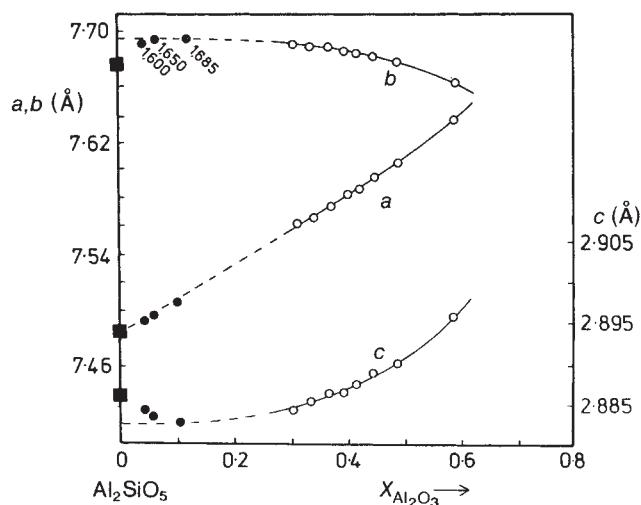


Fig. 2 Lattice parameters of mullites from ref. 4 (open circles) extrapolated to pure Al_2SiO_5 . Filled squares are lattice parameters of sillimanite used as starting material. Filled circles indicate *a*, *b* and *c* parameters for residual sillimanite after annealing at 1,600 °C, 1,650 °C and 1,685 °C. Approximate fits of the experimental data to the extrapolated curves give rough estimates of the compositions of the residual sillimanites. The most aluminous product (~1,685 °C) fits with $X_{\text{Al}_2\text{O}_3} \approx 0.1$ and a composition of $\sim \text{Al}_2[\text{Al}_{2.2}\text{Si}_{1.8}\text{O}_{9.9}]$.

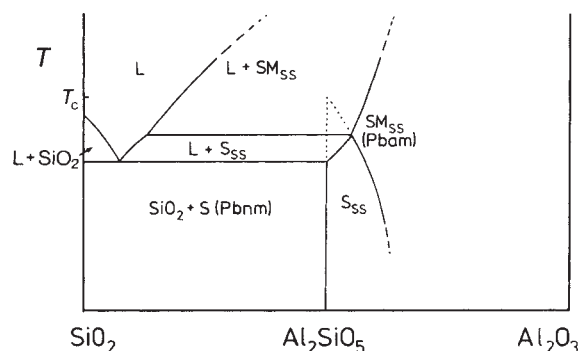


Fig. 3 Postulated phase relations (schematic) of a part of the SiO_2 - Al_2O_3 system close to Al_2SiO_5 for high temperatures at 18–20 kbar. The $\text{Pb}_{\text{nm}} \rightleftharpoons \text{Pb}_{\text{am}}$ transition is shown as having non-first-order character, and a stability field for a sillimanite-mullite solid solution with disordered Al/Si (Pbam) is suggested. L, liquid; S, sillimanite; S_{ss} , sillimanite solid solution; SM_{ss} , sillimanite-mullite solid solution. Dotted lines indicate metastable extensions up to the equilibrium order-disorder transformation temperature (T_c) of pure sillimanite. At lower temperature there is a miscibility gap between sillimanite and mullite^{4,5,22–24}.

The siliceous nature of the melt was suggested initially by the low-pressure melting relations for the system SiO_2 - Al_2O_3 , in which the combination of silica-rich liquid and alumina-rich solid predominates¹⁷. Al-enrichment of the residual sillimanite in our runs is also suggested by changes in lattice parameters (determined from X-ray powder diffraction data) towards those of mullite. Specifically, the sillimanite showed an increase in *a* and *b* dimensions with increasing annealing temperature, while *c* decreased very slightly. Similar changes have been reported in previous high-pressure experiments^{10,11,18,19}, although some authors have attributed them to variations in Al/Si order at

constant composition^{10,11,19}. A very rough estimate of the composition of the final sillimanite after heating at 1,680–1,690 °C can be made by extrapolating the mullite parameters⁴ to pure sillimanite (Fig. 2). Given the short annealing time, this approximate composition does not necessarily represent an equilibrium state.

These experimental observations shed light on the process of mullitization at high pressures, on interpretations of thermochemical data for Al/Si disordering in sillimanite, and on the relationship between sillimanite and mullite. First, when sillimanite is heated at atmospheric pressure it breaks down by a solid-state reaction to mullite + SiO₂ (ref. 6). Our TEM observations suggest that mullitization at 18–20 kbar would occur by partial melting. The total melting interval (from ~1,300 °C to at least ~1,700 °C) is rather large for a simple binary system (Al₂O₃–SiO₂). The apparent melting at 1,300–1,500 °C may depend on the presence of impurities and trace amounts of water, especially since natural sillimanite can contain significant amounts of structural OH groups²⁰. The only impurity detected by electron microprobe analysis of this sample was iron (~0.9 wt % Fe₂O₃, assuming all Fe present as Fe³⁺).

Second, annealing sillimanite at high *P* and *T* induces a significant enthalpy change relative to untreated natural crystals¹¹. Navrotsky *et al.*¹¹ found that the lattice parameters and heats of solution in lead borate of sillimanite crystals heated at 1,400–1,500 °C (16–23 kbar) show relatively small changes, but more substantial effects were found for crystals heated at 1,600–1,700 °C (16–23 kbar). In both temperature ranges, at 18–20 kbar, our annealed products show high dislocation densities, APB-like defects, and inclusions. Some degree of Al/Si disorder at 1,400–1,550 °C may be implied by the calorimetric results and is certainly present at the APB-like defects. Because enthalpies of vitrification are usually substantial compared with enthalpies of ordering²¹, however, it is possible that at least part of the heat-of-solution changes might be attributed to the presence of glass in the annealed products. Moreover, the break in behaviour observed by Navrotsky *et al.* at 1,550–1,600 °C appears to coincide with a change in the proportion of inclusions present in our material. We have examined some of the original samples prepared for the calorimetric study (generously lent to us by Professor R. C. Newton) and, although we were unable to characterize them fully because they were too finely ground to prepare for TEM by ion-beam thinning, they did appear to have similar microstructures to those described above.

By analogy with order-disorder transformations in other mineral solid solutions⁹, the equilibrium order-disorder temperature, *T*_{ord}, for the Pbnm ⇌ Pbam transformation should have a maximum value for pure Al₂SiO₅. At this composition the ratio of tetrahedral Al to Si in sillimanite is 1:1; as the composition deviates from this stoichiometry, *T*_{ord} should decrease. Pbam crystals have been produced in the present experiments at 1,680–1,690 °C by partial melting; that is, they crossed the Pbnm ⇌ Pbam transition by changing composition. *T*_{ord} for pure sillimanite must, therefore, be at a higher temperature (at 18–20 kbar), in agreement with a value of ~1,740 °C estimated on the basis of a statistical mechanical model of the disordering⁸.

Figure 3 shows a schematic phase diagram illustrating the order-disorder and melting relations as interpreted here. Many uncertainties remain but there are some interesting contrasts with solid-solution behaviour at lower pressures. From observations of two-phase assemblages^{4,22} and of sillimanite-mullite intergrowths which developed from initially homogeneous crystals by exsolution^{23,24}, it is clear that there is a miscibility gap between mullite and sillimanite at geological pressures and temperatures^{4,5}. Some natural crystals have been found to have intermediate compositions, but these are ordered, with sillimanite-type ordering reflections^{3,25}. Disordered crystals have not yet been found in nature. We suggest, on the basis of our experiments, that an equilibrium stability field for a Pbam sillimanite–mullite solid solution may exist between ordered sil-

limanite (commensurate, Pbnm) and ordered mullite (incommensurate) at high pressures and temperatures. Further experiments are under way to test this hypothesis and to bracket the order-disorder transformations within the solid solution.

We thank NERC for financial support (grants GR3/4404 to J. D. C. McConnell and GR3/5246 to T.J.B.H. and A. Putnis). This paper is Cambridge Earth Sciences contribution ES 653.

Received 22 October 1985; accepted 22 January 1986.

- Day, H. W. & Kumin, H. J. *Am. J. Sci.* **280**, 265–287 (1980).
- Ribbe, P. H. *Rev. Miner.* **5** (2nd edn), 189–214 (1982).
- Cameron, W. E. *Bull. Am. Ceram. Soc.* **56**, 1003–1007, 1011 (1977).
- Cameron, W. E. *Am. Miner.* **62**, 747–755 (1977).
- Cameron, W. E. *Phys. & Chem. Miner.* **1**, 265–272 (1977).
- Guse, W., Saalfeld, H. & Tjandra, J. *Neues Jb. Miner. Mh.* 175–181 (1979).
- McConnell, J. D. C. & Heine, V. *Phys. Rev. B31*, 6140–6142 (1985).
- Greenwood, H. J. *Mem. geol. Soc. Am.* **132**, 553–571 (1972).
- Carpenter, M. A. *Rev. Miner.* **14**, 187–223 (1985).
- Beger, R. M., Burnham, C. W. & Hays, J. F. *Geol. Soc. Am. Abstr. Prog.* **2**, 490–491 (1970).
- Navrotsky, A., Newton, R. C. & Kleppa, O. J. *Geochim. cosmochim. Acta* **37**, 2497–2508 (1973).
- Burnham, C. W. *Carnegie Inst. Wash. Yb.* **63**, 227–228 (1964).
- Menard, D. & Doukhan, J.-C. *J. Phys. (Fr.)* **39**, L19–L22 (1978).
- Doukhan, J.-C. & Christie, J. M. *Bull. Minér.* **105**, 583–589 (1982).
- Lefebvre, A. & Paquet, J. *Bull. Minér.* **106**, 287–292 (1983).
- Doukhan, J.-C., Doukhan, N., Koch, P. S. & Christie, J. M. *Bull. Minér.* **108**, 81–96 (1985).
- Aksay, I. A. & Pask, J. A. *J. Am. Ceram. Soc.* **58**, 507–512 (1975).
- Hariya, Y., Dollase, W. A. & Kennedy, G. G. *Am. Miner.* **54**, 1419–1441 (1969).
- Chatterjee, N. D. & Schreyer, W. *Contr. Miner. Petrol.* **36**, 49–62 (1972).
- Beran, A., Hafner, St. & Zemann, J. *Neues Jb. Miner. Mh.* 219–229 (1983).
- Navrotsky, A. *Rev. Miner.* **14**, 225–275 (1985).
- Cameron, W. E. *Geol. Mag.* **113**, 497–592 (1976).
- Cameron, W. E. *Nature* **264**, 736–738 (1976).
- Wenk, H.-R. *Neues Jb. Miner. Abh.* **146**, 1–14 (1983).
- Cameron, W. E. *Am. Miner.* **61**, 1025–1026 (1976).

Rapid pressurization experiments on a liquid Al–Mn alloy

J. A. Sekhar & T. Rajasekharan

Defence Metallurgical Research Laboratory,
Hyderabad-500258, India

Rapid pressurization of liquid metals has been shown both theoretically and experimentally to yield appreciable undercooling^{1,2}. With the aim of demonstrating metastable phase formation by this technique, we have now performed a rapid pressurization experiment on an Al₈₀Mn₂₀ alloy and studied the product thus obtained. This was found to contain the metastable icosahedral phase first reported by Shechtman *et al.*³ and studied extensively by many authors^{4–9}. (Pauling⁸ has recently suggested that the same phase can also be indexed to a large cubic cell. However, our conclusions on the formation of metastable phases will be unaffected by the outcome of this controversy.) We have also found that a new metastable cubic ferromagnetic phase of composition Al₃Mn coexists along with the icosahedral phase. The rapid pressurization technique shows promise of becoming a method for obtaining rapidly solidified microstructures and phases in large volumes.

Concomitantly with pressure application on a liquid alloy contained inside a die, three important processes occur: (1) the liquidus rises for all metals that contract on solidification; (2) there is an adiabatic heat release; and (3) the thermal contact at the metal/die interface improves (see ref. 2). If the rise in liquidus temperature exceeds the rise in temperature due to the adiabatic heat, a net undercooling results which can drive rapid solidification. These undercoolings are high, for example >130 K for aluminium when pressurized to 5.7 GPa (which is adequate for complete solidification purely by pressure application). If the adiabatic heat is dissipated by the efficient thermal contact, then the resultant undercooling will be even greater than given above².

Rapid pressurization of the Al₈₀Mn₂₀ alloy was carried out in a simple piston cylinder apparatus made of 'maraging' (high-

Table 1 Observed *d*-spacings and intensities from Al₈₀Mn₂₀ rapidly pressurized alloy

<i>I</i> _{obs}	<i>d</i> (nm)	Icosahedral indices	Observed (<i>d</i> / <i>d</i> (100000) ⁻¹)	Expected (<i>d</i> / <i>d</i> (100000) ⁻¹)	μ phase <i>d</i> (nm)	Intensity	Al ₆ Mn (<i>hkl</i>)	Cubic phase (<i>hkl</i>)	Al ₂ O ₃ (<i>hkl</i>)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
W	0.638				0.636	MW			
VW	0.610				0.607	VW			
VW	0.555				0.549	VW			
VW	0.505				0.505	VW			
W	0.487				0.492	VW	(110)	(111)	
VW	0.438						(002)		
VW	0.413				0.412	VW			
W	0.374	(110001)	0.58	0.56	0.371	W	(020)		
VW	0.331	(111010)	0.65	0.65	0.334	W			
VW	0.295				0.299	MW		(220)	
VVW	0.286						(022)		
W	0.268				0.269	W			
W	0.261						(202)		
St	0.252				0.249	W	(113)	(311)	(104)
W	0.244				0.241	M		(222)	(110)
M	0.232				0.231	MW	(130)		
St	0.226				0.225	MW	(131)		
St	0.215	(100000)	1	1	0.215	M			
St	0.209				0.208	VSt	(310)	(400)	(113)
St	0.204	(110000)	1.058	1.051	0.203	St	(311)		
St	0.201				0.201	M			
MW	0.188				0.189	W	(223)		
VW	0.184						(133)		
VW	0.177				0.177	VW			
VW	0.173						(042)	(422)	(024)
VW	0.169								
VVW	0.164						(224)		
VW	0.161							(333)	
VW	0.160								(116)
VVW	0.153				0.152		(242)		
VVW	0.151								
M	0.148	(111000)	1.45	1.45	0.148	VW	(006)	(440)	
VW	0.147								
VW	0.145	(111100)	1.49	1.49	0.145	VW			
M	0.143				0.143	VW	(243)		
VW	0.140				0.140	M		(531)	(124)
VW	0.137								(030)
W	0.131				0.1306	W			
					0.1296	St			
M	0.1272	(101000)	1.69	1.70	0.1272	St			
M	0.1258				0.1260	W			
					0.1254	W			
W	0.1248				0.1248	VW			
W	0.1238				0.1236	M			
W	0.1226				0.1228	VW			
VW	0.1169	(211000)	1.84	1.82					
VVW	0.1140								
VVW	0.1123				0.1121	MW			
VW	0.1117				0.1115	VW			
VW	0.1093	(110010)	1.97	1.97					
M	0.1070	(200000)	2.01	2.00	0.1070	MW			
					0.1067	MW			
VVW	0.1056				0.1056	W			
VVW	0.1035				0.1036	MW			(226)
VVW	0.1013	(220000)	2.13	2.10	0.1010	VW			
VVW	0.0996								(1,2,10)

A long-exposure (12 h) Debye-Scherrer photograph obtained with Fe K α radiation was used. The icosahedral phase was indexed as explained in the text, the indices for Al₆Mn and α -Al₂O₃ were obtained from the JCPDS¹² powder data file (card nos 06-665 and 10-173) and the *d*-spacings of the μ phase were obtained as described in the text. St, Strong; M, medium; W, weak; V, very.

strength) steel. The alloy was melted in air at 1,245 K in a 3-mm diameter (4-mm height) cylindrical cavity and was rapidly pressurized by forge hammer application on the cold piston. The piston and the liquid alloy were thermally isolated by a hot 1-mm asbestos cloth placed on the liquid metal. We estimate that the sample was pressurized to ~ 2 GPa in < 200 ms (refs 1, 2).

The product of the rapid pressurization experiment was magnetic. A vibrating sample magnetometer measurement of the sample showed the presence of a ferromagnetic phase with a Curie temperature of 800 K. The sample was then analysed by X-ray powder diffraction. The observed *d*-spacings with their assignment to various phases are shown in Table 1. The diffraction pattern from the icosahedral phase (we have used the

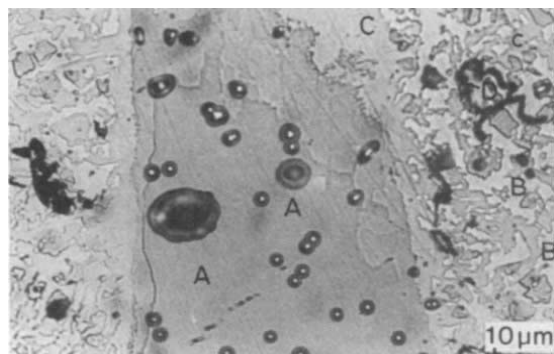


Fig. 1 Microstructure of the rapidly pressurized sample. The compositions of phases A, B, C and D are given in Table 2.

Table 2 Chemical composition (in atomic %) of the various phases shown in Fig. 1

	Al	Mn	Fe	O
Phase A	74.99	24.81	0.20	—
Phase B	78.53	20.58	0.89	—
Phase C	82.41	17.20	0.39	—
Phase D	32.70	1.01	0.09	66.20

icosahedral indexing rather than the cubic indexing of Pauling⁸) was indexed on the basis of six independent vectors $(1, \pm\tau, 0)$, $(0, 1, \pm\tau)$ and $(\pm\tau, 0, 1)$, pointing to the vertices of an icosahedron as done by Bancel *et al.*⁹, where $\tau = 1.618$, the golden mean. In addition to the icosahedral phase, the X-ray powder pattern shows the presence of a cubic phase ($a \approx 0.838$ nm), the μ phase of the Al-Mn system¹⁰, small amounts of Al_6Mn and a little contamination of Al_2O_3 . The X-ray powder pattern after magnetic separation of part of the material showed a clear reduction in intensity of the lines from the cubic phase. The material, when heated at 1,180 K for 3 h in an evacuated quartz ampule (at $\sim 10^{-6}$ torr), was no longer ferromagnetic and X-ray photographs showed a complete transformation of the material to the μ phase (with a little Al_2O_3), showing conclusively that the cubic phase is a new metastable ferromagnetic phase formed by the rapid pressurization process in the Al-Mn system. The transformed phase is the same as that obtained on heating a rapidly quenched $\text{Al}_{78}\text{Mn}_{22}$ alloy at 700 °C for 3 h. A comparison of the X-ray photographs from the rapidly pressurized and rapidly cooled alloys after transformation shows a one-to-one correspondence between the X-ray lines, except for the presence of a few weak lines of Al_2O_3 contaminant in the pressurized sample. Figure 1 shows the typical microstructure of the alloy. The presence of four distinct phases, A, B, C and D, was observed after initially etching mildly with a (FeCl_3 + ethanol + HCl) solution followed by a deep etch with Keller's reagent ($1 \text{ cm}^3 \text{ HF} + 1.5 \text{ cm}^3 \text{ HNO}_3 + 2.5 \text{ cm}^3 \text{ HCl} + 95 \text{ cm}^3 \text{ H}_2\text{O}$). The compositions of these phases were determined using an electron microprobe. Results obtained from the microprobe were corrected for fluorescence, beam absorption and atomic number against pure standards (99.99 wt% pure) and are shown in Table 2. Microprobe results indicate that a small amount of iron from the steel dies had contaminated the material. The dark chain marked D was found to be an Al_2O_3 stringer left over from the melting experiment.

Because the extreme brittleness of the sample precluded transmission electron microscopy, our evidence for correlating the phases to the microstructure is as follows. Phase A could be pitted completely with the FeCl_3 etch when etched for a long time. The other phases were relatively stable to this etch. We melt-spun alloys with compositions from $\text{Al}_{75}\text{Mn}_{25}$ to $\text{Al}_{84}\text{Mn}_{16}$ but could not find any phase that was similarly etched. However, as expected, we found that the icosahedral phases in these

ribbons and those with compositions from $\text{Al}_{84}\text{Mn}_{16}$ to $\text{Al}_{80}\text{Mn}_{20}$ etched similarly to phase C. The pressurization experiments were repeated with a 1-mm die cavity where a higher pressure than that applied earlier could be achieved. The sample thus obtained showed larger areas that would etch with FeCl_3 and also showed higher X-ray intensities of the cubic phase; μ -phase X-ray lines were also then much weaker. We conclude that phase A is the magnetic cubic phase. The μ phase from the equilibrium phase diagram of the Al-Mn system¹⁰ contains ~ 22 atom% Mn, and of phases B and C in Fig. 1, phase B must correspond to the μ phase because of its composition. Phase C is therefore the icosahedral phase. The measured composition of phase C supports this conclusion.

We conclude that the undercooling resulting from rapid pressurization can yield metastable phases, as is shown for the $\text{Al}_{80}\text{Mn}_{20}$ alloy where the icosahedral phase and a new cubic ferromagnetic phase are obtained. The only other ferromagnetic phase in the Al-Mn binary exists at equi-atomic concentration¹¹. The rapid pressurization technique is important because of the possibility of scaling it up to make near net shape components, which have rapidly solidified structures, without relying on rapid heat extraction.

We thank Dr P. Rama Rao for critical discussions and for permission to publish this work.

Received 19 November; accepted 30 December 1985.

1. Sekhar, J. A., Mohan, M., Divakar, C. & Singh, A. K. *Scr. metallurg.* **18**, 1327-1330 (1984).
2. Sekhar, J. A. *Scr. metallurg.* **19**, 1429-1433 (1985).
3. Shechtman, D., Blech, I., Gratias, D. & Cahn, J. W. *Phys. Rev. Lett.* **53**, 1951-xxxx (1984).
4. Levine, D. & Steinhardt, P. J. *Phys. Rev. Lett.* **53**, 2477-XXXX (1984).
5. Maddox, J. *Nature* **314**, 575 (1985).
6. Mackay, A. L. & Kramer, P. *Nature* **316**, 17-18 (1985).
7. Bursill, L. A. & Peng, J. *Nature* **316**, 50-51 (1985).
8. Pauling, L. *Nature* **317**, 512-514 (1985).
9. Bancel, P. A. *et al. Phys. Rev. Lett.* **54**, 2422-2425 (1985).
10. Taylor, M. A. *Acta metall.* **8**, 256-262 (1960).
11. Vlasova, N. I. *et al. Physics Metals Metallogr.*, N.Y. **51** (6), 1-35 (1981).
12. Joint Committee on Powder Diffraction Standards, *Powder Diffraction Data for Metals and Alloys* (International Center for Diffraction Data, Swarthmore, 1978).

Phenol and HCl at 550 °C yield a large variety of chlorinated toxic compounds

Göran Eklund, Jörgen R. Pedersen & Birgitta Strömberg

Studsvik Energiteknik AB, S-611 82 Nyköping, Sweden

During the past decade it has been shown conclusively that the incineration of municipal and industrial wastes gives rise to emissions of chlorinated dibenzodioxins and dibenzofurans. However, the mechanism by which these toxic compounds are formed has not yet been fully established. Some researchers believe that the presence of organically bound chlorine is necessary¹, but others consider that inorganic forms of chlorine may also participate in the process². We now report the synthesis of a large number of chlorinated environmental pollutants in a simple high-temperature experiment. The results show that phenol and HCl are the most likely precursors of the chlorinated dibenzodioxins and dibenzofurans formed in the combustion of wastes. The dependence of the reaction on the concentration of HCl indicates a way of controlling the formation of these toxic compounds during incineration.

Phenol (1 mg) was mixed in quartz tubes with an aqueous solution (10 μl) containing a constant amount of Cl^- (3.7 mg), and variable $[\text{H}^+]$ ($\text{pH} = 0.75-14$). The tubes were sealed and heated to 550 °C for 5 min. After cooling, the residues were extracted with 100 μl of *n*-hexane. The extracts were analysed by gas chromatography-mass spectroscopy (GC-MS) and gas chromatography with electron capture detection. Aliquots (2 μl) of each extract were injected in splitless mode into a Hewlett-Packard 5880 A gas chromatograph with a fused silica column

Table 1 Chlorinated organic compounds produced by reaction of phenol with HCl

GC separate no.	MS identification	No. of isomers	Relative concentration	GC separate no.	MS identification	No. of isomers	Relative concentration
1	Monochlorobenzene	2	m	29	Tetrachlorochromone	1	t
2	Chlorophenol	2	m	30	Dichlorodibenzofuran	2	m
3	Dichlorobenzene	3	m	31	Trichloronaphthoquinone	1	t
4	Tetrachloropropene	1	m	32	Tetrachloronaphthoquinone	1	t
5	Pentachloropropene	2	m	33	Monochloroxanthrone	1	t
6	Trichlorobenzene	3	M	34	Pentachlorodihydroxy- cyclohexane	2	m
7	Dichlorophenol	2	M	35	Trichlorodibenzofuran	5	m
8	Dichlorobenzofuran	1	M	36	Trichlorodibenzodioxin	5	m
9	Tetrachlorobenzene	3	M	37	Dichloroxanthrone	1	t
10	Trichlorophenol	2	M	38	Tetrachlorodibenzofuran	12	t
11	Trichlorocresol	1	m	39	Tetrachlorodibenzodioxin	>10	m
12	Dichloroindenone	3	t	40	Trichlorobiphenyl	1	t
13	Trichlorobenzofuran	2	m	41	Trichloroxanthrone	1	t
14	Pentachlorobenzene	1	m	42	Pentachlorobiphenyl	2	m
15	Tetrachlorophenol	2	M	43	Pentachlorochromone	2	t
16	Tetrachlorocresol	1	m	44	Pentachlorodibenzofuran	5	t
17	Tetrachlorobenzofuran	3	m	45	Pentachlorodibenzodioxin	5	t
18	Trichloroindenone	3	m	46	Tetrachloroxanthrone	1	m
19	Trichloromethyl- benzoquinone	1	m	47	Hexachlorocoumarone	1	t
20	Hexachlorobenzene	1	m	48	Hexachlorodibenzodioxin	5	m
21	Monochlorodibenzofuran	2	t	49	Hexachlorodibenzofuran	4	m
22	Dichloronaphthoquinone	1	t	50	Tetrachlorodihydroxy- tetrahydrophenanthrenone	1	t
23	Pentachlorophenol	1	m	51	Heptachlorodibenzofuran	1	t
24	Tetrachlorobenzodi- hydroquinone	1	m	52	Dichlorotriphenylbenzene	1	t
25	Dichlorohydroxyindenone	1	t	53	Octachlorodibenzodioxin	1	t
26	Dichlorochromone	1	t	54	Octachlorodibenzofuran	1	t
27	Tetrachlorobenzoquinone	2	m	55	Hexachlorotriphenylbenzene	1	t
28	Trichlorochromone	1	t				

M, m and t denote major, minor and trace reaction products, respectively.

(SE-54, 30 m, 0.22 mm i.d.). The GC-MS analyses were performed on a Finnigan 1020 GCMS, and the mass spectra were tentatively identified by comparison with standard mass spectra libraries. Table 1 contains the result of the GC-MS analysis of the extract from the experiment at highest [HCl]; the list of identified chlorinated organic compounds is very similar to lists of pollutants found in flue gases from municipal incinerators.

Figure 1 shows the amount of phenols produced, as a function of [HCl]; clearly, the reaction is very sensitive to [HCl]. At [HCl] $< 10^{-3}$ M no chlorinated compounds could be detected.

Although the reaction time in these experiments was substantially longer than the residence time of fuels during incineration, the temperatures and HCl concentrations were such as may be achieved locally in a typical waste incinerator. Phenol and precursors to phenol are ubiquitous in household wastes. A reasonable hypothesis is that chlorinated organic compounds are produced from phenols, acids and any chlorine source in

the combustion zone. This hypothesis can easily be tested in a full-scale experiment in which acids from the fuels are neutralized during combustion.

We believe that production of chlorinated dibenzodioxins and dibenzofurans is a probable, though not a necessary result of waste incineration. The high sensitivity of the chlorination reaction to [HCl] suggests that the emissions of these pollutants can be significantly reduced at low cost. Reduction of [HCl] in the combustion zone should be possible by injection of a suitable alkaline sorbent along with the fuel.

Received 28 October 1985; accepted 24 January 1986.

1. Rappe, Ch., Markland, S., Buser, H. R. & Bosshard, H. P. *Chemosphere* 7, 269-277 (1978).
2. Bumb, R. R. et al. *Science* 210, 385-390 (1980).

Measurement of organic carbon in polar snow samples

Mark S. Twickler, Mary Jo Spencer,
W. Berry Lyons & Paul A. Mayewski

Glacier Research Group and Department of Earth Sciences,
University of New Hampshire, Durham,
New Hampshire 03824, USA

Glaciers provide a unique medium for the study of palaeoatmospheric chemistry¹⁻³, particularly in polar glaciers where chemical records may extend for at least several hundred thousand years. Glaciochemical records from remote, high-latitude areas provide a regional- to global-scale integration of past atmospheric chemistry. These records yield information regarding climate change⁴⁻⁷, volcanic events^{8,9} and solar activity¹⁰, as well as the effect of human activities^{11,12}. Although gaseous forms of organic carbon have been analysed in polar ice^{13,14}, we present here the first measurements, to our knowledge, of organic carbon from

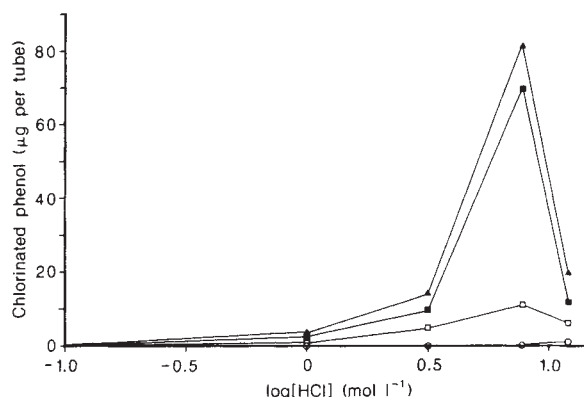


Fig. 1 The formation of chlorinated phenols as a function of HCl concentration. At [HCl] lower than 10^{-3} M, no chlorinated compounds could be detected. ○, Pentachlorophenol; ■, tetrachlorophenols; □, trichlorophenols; ▲, total chlorophenols. Reaction conditions are specified in the text.

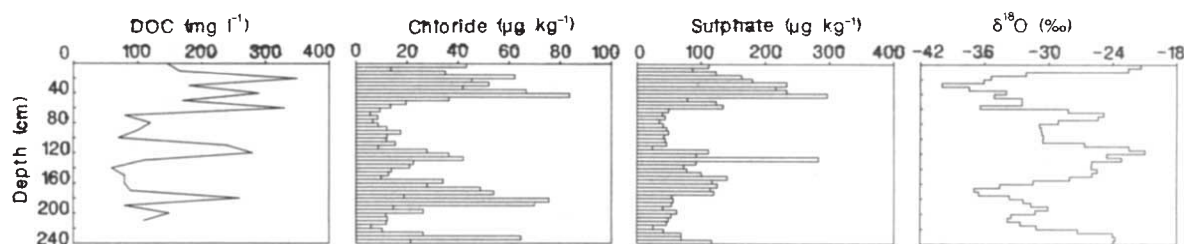


Fig. 1 Concentrations of DOC, chloride, sulphate and $\delta^{18}\text{O}$ in snowpit.

polar firn samples. Our measurements are among the lowest reported for organic carbon in precipitation. The data indicate that dissolved organic carbon in Greenland snow has a seasonal deposition pattern, with higher concentrations observed in the winter/spring period. Although we are unable to establish the source of the organic carbon in Greenland snow, three possibilities are discussed.

The burning of fossil fuels and the clearing of forests have led to the increase in inorganic carbon input into the atmosphere that has been documented recently in polar ice cores¹⁵. These processes introduce organic carbon and elemental carbon into the atmosphere as well^{16,17}. It is also evident that natural processes may introduce quantities of gases and particulate organic carbon into the atmosphere¹⁸⁻²³. We now compare organic carbon data for snow from a remote, high-latitude location with values from other, lower latitude locations and evaluate the feasibility of using ice core samples to monitor the historic record of both natural and anthropogenic organic carbon emissions into the atmosphere.

In June 1984, a 2.3-m snowpit was carefully excavated ~40 km south-west of Dye 3, Greenland (65.01° N, 44.87° W). The snowpit sampling was part of a larger programme that involved the electrochemical drilling of a 115-m core at this location and the collection of snow samples from several 1.5-m snowpits between the drilling site and Dye 3. The sampling was carefully undertaken by individuals wearing polyethylene gloves, non-particulating clean-suits and particle masks. Samples for organic carbon analysis were collected at 10-cm intervals by pushing Pyrex tubes (previously combusted at 500 °C), horizontally into a cleanly scraped side of the snowpit. A tube was taken into the field, opened and closed and returned to our laboratory, filled with organic-carbon-free water and analysed as a blank. This blank was below our detection limit of 0.01 mg l⁻¹. The mean of organic carbon values from 19 snow samples collected in snowpits from the depth range 0–150 cm between the primary sampling site and Dye 3 was 0.11 mg l⁻¹ with a range between 0.03 and 0.32 mg l⁻¹. All samples were kept frozen until a few hours before analysis. Samples were analysed by injecting 10 cm³ of melted snow into a Barnstead Photochem organic carbon analyser with a glass syringe. Potassium hydrogen phthalate was used as a standard. All glassware and syringes were cleaned and/or rinsed several times in Chromerge before use. Duplicate analyses of four samples gave the following results: 0.145 ± 0.005, 0.165 ± 0.055, 0.10 ± 0.01 and 0.07 ± 0.01 mg l⁻¹.

We report the values as dissolved organic carbon (DOC) even though the samples were not filtered before analysis. A few minutes before the samples were extracted from their containers, the tubes were swirled in order to homogenize the samples; this may have resuspended any particulate matter present. Samples were then drawn from the top of the liquid. Although colloidal organic material present in the samples was undoubtedly collected and analysed by this technique, much of the particulate carbon present in these samples remained in the sample containers. Recent work at Dye 3 indicates that only a very small percentage of the tropospheric aerosol collected in the spring season is composed of organic carbon²⁴. In addition, owing to the nature of our analytical technique (ultraviolet oxidation of

the sample), it is highly unlikely that the measurement includes any elemental carbon (soot) present in the sample. No attempt was made to assess possible adsorption of DOC onto the container walls during the thawing procedure.

Figure 1 shows a plot of the DOC concentration versus depth in the snowpit. In addition, the values for Cl⁻ and SO₄²⁻ (measured by ion chromatography) are also shown. The DOC values are the lowest ever reported for precipitation (Table 1). In general, the DOC profile resembles those of Cl⁻ and SO₄²⁻. Busenberg and Langway have demonstrated that the highest Cl⁻ values in Dye 3 precipitation occur in the winter/spring period²⁵. Examination of the $\delta^{18}\text{O}$ data from this snowpit (Fig. 1) suggest that the high DOC values are generally associated with lower temperatures (the correlation between Cl⁻ and $\delta^{18}\text{O}$ is negative and significant at the 95% confidence level). Because the major source of Cl⁻ to this area is thought to be sea-salt aerosol^{9,25}, the DOC probably has a similar source, or is transported along the same pathway.

One explanation for the increased marine aerosol component in the winter/spring snows of south-central Greenland is that it reflects increased storm activity, wave action and sea-salt entrainment in the ocean and atmosphere surrounding Greenland at this time of year. 'Clean' marine aerosols at ~41° S latitude contain a considerable amount of organic matter²⁶. Another explanation of this seasonal pattern of DOC deposition is that the sources of these chemical species (Cl⁻, SO₄²⁻ and DOC) may be distinct, but common atmospheric transport and removal processes lead to a similarity in depositional pattern. For example, the continentally derived element aluminium also shows a strong late winter/spring peak in central Greenland snow²⁷ similar to that of marine-derived species such as Cl⁻. Therefore, a long-travelled continental biogenic source of the DOC is also possible. Andreae has shown that most organic carbon present in the tropospheric aerosol over the Atlantic Ocean has a terrestrial source¹⁶. These compounds could be the products of gas-to-particle conversion of continentally derived gaseous precursors²⁸. Finally, a local to regional terrestrial source of the DOC is highly unlikely as higher values would be expected primarily during the summer²⁹. High SO₄²⁻

Table 1 Mean dissolved organic carbon values for remote precipitation

Location	Concentration of DOC (mg l ⁻¹)	Ref.
Carrol Glacier, Alaska	0.15	31
Hubbard Brook, New Hampshire	1.09	32
Ithaca, New York	1.88	27
Como Creek, Colorado	1.52	33
Enewetak Atoll	0.27 (rainy season) 1.18 (dry season)	34
Alberta	1.6	35
Colorado Rockies†	0.5	36
Amsterdam Island	As high as 0.4*	23
Katherine, Australia	As high as 0.6*	23

* Values only for acetic and formic acid carbon.

† Subsurface values.

concentrations in Arctic aerosols and snows during the winter/spring period are due to meteorological conditions that allow transport of air masses from anthropogenically impacted mid-latitude areas to the Arctic³⁰. The observed winter/spring DOC maxima may then also represent deposition of a continentally derived aerosol which has been affected by fossil fuel burning.

Obviously, we lack the data to determine the origin of the DOC at this time and we cannot rule out the possibilities that the DOC may: (1) reflect an input of marine organic carbon and be similar in source to Cl^- ; (2) represent a natural continental source similar to Al; or (3) have its origin from an anthropogenic source similar to the winter/spring SO_4^{2-} signal. Only by the detailed investigation of the specific organic compounds present in polar snows will a unique solution to the origin of DOC be found. It does appear, however, no matter what the source, that DOC may be useful as a seasonal indicator in Greenland ice.

Owing to the small number of samples analysed, the short time period represented by the snowpit and the lack of identification of specific organic compounds, we cannot comment with any certainty on the effect of human activities on the DOC concentrations found in the Greenland snow. More detailed studies using ice-core samples could provide detailed information regarding historic organic carbon emissions as well as the ultimate source of the DOC.

This work was supported by a contract from the US Environmental Protection Agency administered through North Carolina State University (APP-0306-1983). We thank W. Dansgaard and H. Clausen for the oxygen isotope analyses and J. V. James and T. Hinkley for help with sample collection.

Received 19 August; accepted 30 December 1985.

1. Boudron, C. & Lorius, C. *Ambio* **9**, 210-215 (1980).
2. Herron, M. M. & Herron, S. L. *J. Sci. Hydrol.* **28**, 139-153 (1983).
3. Wolff, E. W. & Peel, D. A. *Nature* **313**, 535-540 (1985).
4. Thompson, L. G. & Mosley-Thompson, E. *Science* **212**, 812-815 (1981).
5. Mayewski, P. A. & Lyons, W. B. *Geophys. Res. Lett.* **9**, 190-192 (1982).
6. Palais, J. M. & Legrand, M. *J. geophys. Res.* **90**, 1143-1154 (1985).
7. Finkel, R. C. & Langway, C. C. Jr *Earth planet. Sci. Lett.* **73**, 196-206 (1985).
8. Hammer, C. U., Clausen, H. B. & Dansgaard, W. *Nature* **288**, 230-235 (1980).
9. Herron, M. M. *J. geophys. Res.* **87**, 3052-3060 (1982).
10. Raisbeck, G. M. *et al.* *Nature* **293**, 825-826 (1982).
11. Ng, A. & Patterson, C. *Geochim. cosmochim. Acta* **45**, 2109-2121 (1981).
12. Boudron, C. F. & Patterson, C. C. *Geochim. cosmochim. Acta* **47**, 1355-1368 (1983).
13. Craig, H. & Chou, C. C. *Geophys. Res. Lett.* **9**, 1221-1224 (1982).
14. Rasmussen, R. A. & Khalil, M. A. K. *J. geophys. Res.* **89**, 11599-11605 (1984).
15. Nefel, A., Moor, E., Oeschger, H. & Stauffer, B. *Nature* **315**, 45-47 (1985).
16. Andreae, M. O. *Science* **220**, 1148-1151 (1983).
17. Hanse, A. D. A. & Rosen, H. *Geophys. Res. Lett.* **11**, 381-384 (1984).
18. Rasmussen, R. A. & Went, F. W. *Proc. natn. Acad. Sci. U.S.A.* **53**, 215-220 (1956).
19. Lewis, W. M. Jr *Wat. Resour. Res.* **17**, 169-181 (1981).
20. Andreae, M. O. & Raemdonck, H. *Science* **221**, 744-777 (1983).
21. Cline, J. D. & Bates, T. S. *Geophys. Res. Lett.* **10**, 949-952 (1983).
22. Hov, Ø, Penkett, S. A., Isaksen, I. S. A. & Semb, A. *Geophys. Res. Lett.* **11**, 425-428 (1984).
23. Galloway, J. M., Likens, G. E., Keene, W. C. & Miller, J. M. *J. geophys. Res.* **87**, 8771-8786 (1982).
24. Davidson, C. I., Santhanam, S., Fortmann, R. C. & Olson, M. P. *Atmos. Envir.* (submitted).
25. Busenberg, E. & Langway, C. C. Jr *J. geophys. Res.* **84**, 1705-1709 (1979).
26. Andreae, M. O. *J. geophys. Res.* **87**, 8875-8885 (1982).
27. Langway, C. C. Jr, Cragin, J. H., Klouda, G. A. & Herron, M. M. *Proc. IUGG Symp. Isotopes and Impurities in Snow and Ice*, 302-306 (International Association of Hydrological Sciences, 1977).
28. Chesselet, R., Fontugne, M., Buat-Menard, P., Ezat, U. & Lambert, C. E. *Geophys. Res. Lett.* **8**, 345-348 (1981).
29. Likens, G. E., Edgerton, E. S. & Galloway, J. M. *Tellus* **35B**, 16-24 (1983).
30. Rahn, K. A. *Atmos. Environ.* **15**, 1447-1455 (1981).
31. Loder, T. C. & Hood, D. W. *Limnol. Oceanogr.* **17**, 349-355 (1972).
32. Likens, G. E., Bormann, F. H., Pierce, R. S., Eaton, J. S. & Johnson, N. M. *Biogeochemistry of a Forested Ecosystem* (Springer, New York, 1977).
33. Grant, N. C. & Lewis, W. M. Jr *J. Tellus* **34**, 74-88 (1982).
34. Zafriou, O. V., Gagosian, R. B., Peltzer, E. T., Alford, J. B. & Loder, T. C. *J. geophys. Res.* **90**, 2409-2423 (1985).
35. Wallis, P. M., Hynes, H. B. N. & Telang, S. A. *Hydrobiologia* **79**, 77-90 (1981).
36. Thurman, E. M. *Organic Geochemistry of Natural Waters* (Nijhoff/Junk, Dordrecht, 1985).

Significance of atmospheric-derived fixed nitrogen on productivity of the Sargasso Sea

Anthony Knap & Timothy Jickells*

Marine and Atmospheric Programme, Bermuda Biological Station for Research, Inc., Ferry Reach 1-15, Bermuda

Alex Pszeny & James Galloway

Department of Environmental Studies, University of Virginia, Charlottesville, Virginia 22903, USA

Recently, it has been suggested that marine primary productivity can be influenced by direct inputs of fixed inorganic nitrogen (NO_3^- , NO_2^- and NH_4^+) from atmospheric deposition, and that this may be important in shaping trends of productivity in coastal areas¹. It is known that anthropogenic contaminants are transported from North America to the Sargasso Sea, affecting the chemistry of rain water over this oligotrophic ocean area. Here we present data collected between 1982 and 1984 as part of the Western Atlantic Ocean Experiment (WATOX) by a wet-only precipitation collector on Bermuda. Our measurements of the anthropogenic fixed nitrogen in the rainwater samples are discussed in the context of the amount of nitrogen required to support 'new' primary production in the Sargasso Sea. We show that the effects of atmospheric deposition are insignificant.

The suggested influence of precipitation on primary productivity was derived using sea water diluted with nutrient-rich deionized water and/or one or two natural local rainwater samples of unspecified composition. As this suggestion has severe implications for oceanic nutrient budgets, we felt it

required further consideration, especially as the nitrogen content of rain water has increased over the past 100 years because of anthropogenic emissions².

We have chosen the Sargasso Sea as a useful area in which to evaluate the effects of rain-derived fixed nitrogen sources for four main reasons:

- (1) It is a central ocean gyre, removed from coastal sources, with low rates of primary productivity³ and therefore potentially susceptible to.
- (2) It has been shown to be 'downwind' of the North American continent which is a source of acid rain containing fixed nitrogen from anthropogenic emissions; long-range transport of this material can deposit this anthropogenic fixed nitrogen in the Sargasso Sea⁴.
- (3) There is a large database available on the precipitation chemistry on Bermuda which can be extrapolated to estimate wet deposition in the Sargasso Sea^{5,6}. There is also a considerable database of precipitation events monitored by ships in the Sargasso Sea⁷.
- (4) There is a long-term (31-yr) database on the hydrography of this area⁸ and considerable data on primary productivity collected periodically at 32°N, 65°W over the same time period^{3,9}.

We have therefore attempted to assess the relative significance of the atmospheric fixed nitrogen source on the primary productivity of the Sargasso Sea using the considerable data available.

From September 1982 to May 1984 we collected rain water on an event basis at High Point, Bermuda. Samples were spiked immediately after collection with a biocide (CHCl_3) to prevent chemical changes. Samples were analysed at the University of Virginia for NO_3^- and NH_4^+ ions using colorimetric and ion chromatographic techniques¹⁰. The results for 126 samples are summarized in Table 1 as deposition rates, in $\mu\text{mol N per m}^2$ per event, calculated from the measured NH_4^+ and NO_3^- concentrations and the amount of precipitation for each event¹¹. There is negligible NO_2^- in Bermuda rain^{4,5}. Ranges and

* Present address: School of Environmental Science, University of East Anglia, Norwich NR4 7TJ, UK.

Table 1 Wet deposition of $\text{NH}_4^+\text{-N}$ and $\text{NO}_3^-\text{-N}$ at High Point, Bermuda from September 1982 to May 1984

Component	Minimum	Geometric mean ($\mu\text{mol N per m}^2$ per event)	Maximum	Volume-weighted average* ($\mu\text{mol N m}^{-2} \text{ d}^{-1}$)
$\text{NH}_4^+\text{-N}$	1	32	470	13
$\text{NO}_3^-\text{-N}$	3	42	500	17
$(\text{NH}_4^+ + \text{NO}_3^-)\text{-N}^\dagger$	7	77	860	30

For details of analyses, see ref. 10.

* Calculated as volume-weighted average of $\text{NH}_4^+\text{-N}$ or $\text{NO}_3^-\text{-N}$ multiplied by the average precipitation in Bermuda of 152.5 cm yr^{-1} (ref. 4), then divided by 365.25.

† These values are different from the sum of NH_4^+ and NO_3^- because the minimum and maximum values of NH_4^+ and NO_3^- occurred in different events, thus affecting the totals.

geometric means are listed as well as the deposition rate in $\mu\text{mol m}^{-2} \text{ day}^{-1}$, calculated from the volume-weighted average and the average annual precipitation for Bermuda (152.5 cm^4). The NO_3^- appears to have a predominantly anthropogenic source in Bermuda precipitation⁴ and can therefore be assumed to have increased over the past 100 yr because of increased anthropogenic emissions² from North America. Sources of ammonium are probably related to agricultural practices in North America² and appear to travel with the acids in Sargasso Sea precipitation^{3,7}. There is little information on the concentrations of organic nitrogen in rain, its form or whether it can be used for primary productivity. The organic nitrogen values that are available represent total digestions and are therefore overestimates¹². The concentration of organic nitrogen has been calculated to be $7.5 \mu\text{mol l}^{-1}$, which corresponds to a flux of $31.2 \mu\text{mol m}^{-2} \text{ d}^{-1}$, or approximately equal to the NH_4^+ and NO_3^- input.

Estimates of total primary production rates in the Sargasso Sea have been calculated as $72 \text{ g C m}^{-2} \text{ yr}^{-1}$ (ref. 3), and more recent measurements by our group yielded a similar value of

$90 \text{ g C m}^{-2} \text{ yr}^{-1}$ (S. R. Smith, T.J. and A.K., unpublished data). Using the 'Redfield ratio'¹³ of C/N of 106:16 or the more recent ratio of 122:16 (ref. 14), and assuming that 10–20% of net productivity is new production¹⁵, there is a nutrient requirement of $\sim 0.08\text{--}0.23 \text{ mol N m}^{-2} \text{ yr}^{-1}$ or $220\text{--}630 \mu\text{mol N m}^{-2} \text{ d}^{-1}$. The rates of primary productivity in the Sargasso Sea have recently been critically reviewed⁹; based on oxygen budget analysis, it is estimated that new (not 'net') productivity may be far greater than the proposed 10–20% of the net productivity rates¹⁵. It is probable that new productivity is as high as $50 \text{ g C m}^{-2} \text{ yr}^{-1}$ (ref. 9) which would then require $1,700 \mu\text{mol N m}^{-2} \text{ d}^{-1}$.

We have created an inventory of nitrogen sources to the Sargasso Sea based on our work and that of others (Table 2). Even with the inclusion of organic nitrogen sources, nitrogen input due to precipitation ranges from 16 to $80 \mu\text{mol N m}^{-2} \text{ d}^{-1}$. The 'low' of $16 \mu\text{mol N m}^{-2} \text{ d}^{-1}$ is based on the possibility that organic nitrogen is not a usable source of nitrogen.

This 'new' production may be supported by nutrient injection from below the euphotic zone. Klein and Coste's 1984 model¹⁶ suggests that nutrients are injected in pulses with frequencies in hours in response to wind stress interaction with surface currents. Klein and Coste's average value for an oligotrophic area in the Mediterranean Sea is $\sim 900 \text{ mmol N m}^{-2} \text{ yr}^{-1}$, which corresponds to an average of $2,500 \mu\text{mol N m}^{-2} \text{ d}^{-1}$. Although we cannot extrapolate this result to the oligotrophic Sargasso Sea, it does serve to illustrate the range and magnitude of fixed nitrogen that could be provided by this source. Even if there was one major precipitation event, and assuming this lasted 1 day and contained the maximum amount of fixed nitrogen that we have found (Table 1), it could supply only half the nitrogen required for new production⁹.

As the Sargasso Sea is, in our opinion, an extreme case because of the significant anthropogenic component in North American rain, and because this anthropogenic component only reaches Bermuda during favourable meteorological events, we feel we can generalize about other marine oligotrophic areas. We therefore conclude that precipitation plays a minor part in the provision of nutrients for primary productivity in oligotrophic ocean areas^{17,18} even though there has been a considerable increase in nitrogen in rain water caused by anthropogenic sources over the past 100 years.

We thank J. Tokos and W. Keene for collection and analysis of the rain water, S. R. Smith for the primary productivity data, and J. Goldman and R. Duce for valuable comments on the manuscript. This work was supported by the US NOAA under the WATOX funding. Additional support was provided by the Bermuda Government. This is contribution no. 1081 from the Bermuda Biological Station for Research, Inc.

Received 11 October 1985; accepted 15 January 1986.

1. Paerl, H. W. *Nature* **315**, 747–749 (1985).
2. Brimblecombe, P. & Stedman, D. H. *Nature* **298**, 460–462 (1982).
3. Menzel, D. & Ryther, J. *Deep-Sea Res.* **6**, 351–367 (1960).
4. Jickells, T. D., Knap, A. H., Church, T., Galloway, J. N. & Miller, J. *Nature* **297**, 55–57 (1982).
5. Church, T. M., Galloway, J. N., Jickells, T. D. & Knap, A. H. *J. geophys. Res.* **87**, 11013–11018 (1982).

* Estimates of dry deposition rates to the Sargasso Sea have not been made; these represent estimates of global averages.

† Recalculated based on ref. 19. (Originally calculated as $225\text{--}1,350 \mu\text{mol N m}^{-2} \text{ d}^{-1}$, however a mathematical error in conversion of nitrate flux to $\text{m}^{-2} \text{ h}^{-1}$ resulted in an order-of-magnitude overestimate.)

‡ S. R. Smith, T.J. and A.K., unpublished data.

6. Galloway, J. N. in *The Biogeochemical Cycling of Sulfur and Nitrogen in the Remote Atmosphere* (eds Galloway, J. N., Charlson, R. J., Andrae, M. O. & Rodhe, H.) (Reidel, Dordrecht, in the press).
7. Galloway, J. N., Knap, A. H. & Church, T. M. *J. geophys. Res.* **88**, 10859–10864 (1983).
8. Wright, W. R. & Knap, A. H. *Intergovt oceanogr. Commn Tech. Ser.* **24** (1983).
9. Jenkins, W. J. & Goldman, J. C. *J. mar. Res.* **43**, 465–491 (1985).
10. Galloway, J. N., Likens, G. E., Keene, W. C. & Miller, J. J. *J. geophys. Res.* **87**, 8771–8786 (1982).
11. Keene, W. C., Pszeny, A. A. P., Galloway, J. N. & Hawley, M. E. *J. geophys. Res.* (submitted).
12. Jickells, T. D. & Hillier, G. B. in *Bermuda biol. Stn Spec. Publ.* No. 18, 63–66 (1982).
13. Redfield, A. C., Ketchum, B. H. & Richards, F. A. in *The Sea* Vol. 2 (ed. Hill, M. N.) (Interscience, London, 1963).
14. Takahashi, T., Broecker, W. S. & Langer, S. J. *J. geophys. Res.* **90**, 6907–6924 (1985).
15. Dugdale, R. C. & Goering, J. J. *Limnol. Oceanogr.* **12**, 196–206 (1967).
16. Klein, P. & Coste, B. *Deep Sea Res.* **31**, 21–37 (1984).
17. Carpenter, E. D. & McCarthy, J. J. *Limnol. Oceanogr.* **20**, 389–401 (1975).
18. Menzel, D. W. & Spaeth, J. P. *Limnol. Oceanogr.* **7**, 159–162 (1962).
19. McCarthy, J. J. & Carpenter, E. D. *Nitrogen in the Marine Environment* (eds Carpenter, E. J. & Capone, D. G.) 487–512 (Academic, New York, 1983).

Isoprenoid thiophenes: novel products of sediment diagenesis?

S. C. Brassell*, C. A. Lewis*, J. W. de Leeuw†, F. de Lange† & J. S. Sinninghe Damsté†

* Organic Geochemistry Unit, University of Bristol, School of Chemistry, Cantock's Close, Bristol BS8 1TS, UK
 † Organic Geochemistry Unit, Delft University of Technology, Department of Chemistry and Chemical Engineering, de Vries van Heystplantsoen 2, 2628 RZ Delft, The Netherlands

Sulphur is a significant component of the organic matter in recent and ancient sediments and in petroleum^{1,2}, yet the precise nature of its association and incorporation is poorly understood. Various sulphur-containing compounds have been recognized in petroleum^{2–4}, but little is known about their origins and mode of generation during sediment burial, and for only a few organo-sulphur compounds with >15 carbon atoms have the structures been determined^{5,6}. Here we identify one of the alkyl thiophenes which occur widely in both recent and ancient deep-sea sediments^{7–13} as 3-methyl-2-(3,7,11-trimethyldodecyl)-thiophene, occurring as a limited number of the possible stereoisomers. This compound is presumed to originate from the incorporation of sulphur into chlorophyll-derived phytol, or archaeobacterial

phytenes or their diagenetic products. Its recognition suggests a novel diagenetic pathway for acyclic isoprenoids involving the introduction of sulphur into specific lipid moieties. Similar, but intermolecular, sulphur incorporation might give rise to sulphur-linked macromolecular materials and thereby contribute significantly to the formation of kerogens.

Significant developments have recently occurred in the understanding of the biological origins of sedimentary acyclic isoprenoids and their subsequent diagenetic fate. It is now evident that such compounds, which were among the first clearly related to natural products to be recognized in sediments and petroleum^{14,15}, are not wholly derived from the phytyl side-chain of chlorophyll, but may also originate from the free and ether-bound lipids of archaeobacteria^{16–18} and from the tocopherols¹⁹ of photosynthetic organisms. Here we further expand the range of such compounds with the characterization of acyclic isoprenoid thiophenes in sediments.

In addition to a thiophene-containing hopane⁶, many immature sediments, notably those recovered by the Deep Sea Drilling Project (DSDP), contain a number of components tentatively identified as thiophenes from their mass spectra^{7–10} (Table 1, Fig. 1). These spectra are dominated by cleavages and rearrangements associated with the thiophene ring^{20,21}, a feature that simplifies the recognition of the size of their alkyl substituents, although providing no indication of the structure of these alkyl groups (whether straight-chain or branched). The widespread occurrence and relative abundance of the C₂₀ alkyl thiophenes I and II suggested to us that these compounds might be related to acyclic isoprenoids. Certainly their occurrence in several sediments (for example, in the Japan Trench, Walvis Ridge and Cariaco Trench) with high concentrations of phytenes²² and phytol²³ provided circumstantial evidence for this relationship.

If I (Table 1, Fig. 1) is an acyclic isoprenoid thiophene, then its mass spectrum⁹ might correspond to that expected for 3-methyl-2-(3,7,11-trimethyldodecyl)-thiophene (Fig. 2). Synthesis of this compound (together with its 4-methyl isomer, III; Fig. 3) followed by gas chromatograph (GC) and gas chromatograph-mass spectrometer (GC-MS) coinjection (OV-1 methyl silicone fluid columns) confirmed this assignment. Such a structure might arise from the incorporation of sulphur into a phytadiene or phytol, a process also observed in heating experiments with H₂S and carbohydrates²⁴. The recognition of alkyl

Table 1 Occurrence of alkylthiophenes in oceanic and other sediments

Location	DSDP Leg-Site	Age	Compounds			Ref.
			I	II	Others*	
Cariaco Trench	15–147	Quaternary	✓	✓	—	‡
Japan Trench	56–436	Pliocene	—	—	F	‡
	57–440	Pleist.-Mio.	✓	✓	E, F	‡
San Miguel Gap	63–467	Plio.-Mio.	✓	✓	F	7
Gulf of California	64–474	Pleistocene	—	—	F	8
	64–478	Pleistocene	—	—	G, H	8
	64–479	Pleistocene	✓	✓	G, H	‡
	64–481	Pleistocene	—	—	Unspecified†	8
Middle America Trench	67–496	Quaternary	✓	✓	—	‡
Walvis Ridge	75–532	Quat.-Plio.	✓	✓	A–F	9,‡
Angola Basin	75–530	Miocene	✓	✓	F	9
Mazagan Escarpment	79–545	Cenomanian	✓	—	—	10
	79–547	Eocene	✓	✓	—	10
Livello Bonarelli	NA	Cen./Tur.	✓	✓	—	11
Namibian Shelf	NA	Quaternary	✓	✓	—	12
Sarcina	NA	Miocene	✓	✓	B, C and others	13

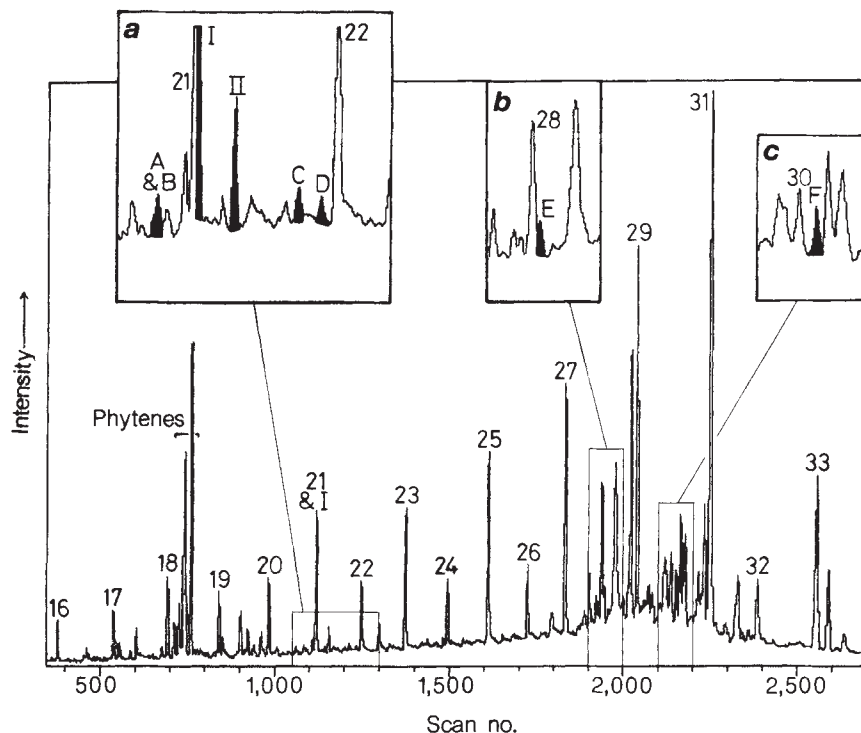
—, Not observed or reported; NA, not applicable. Pleist., Pleistocene; Mio., Miocene; Quat., Quaternary; Plio., Pliocene; Cen., Cenomanian; Tur., Turonian.

* The identities of compounds A to F are given in the legend of Fig. 1; G and H are isomeric C₂₅ alkylthiophenes (M⁺ = 378, prominent ions with mass/charge ratios *m/z* = 265 and 125 (ref. 8)) of unknown structure.

† Range of mono-, di- and trialkylated thiophenes.

‡ S.C.B., unpublished data.

Fig. 1 Reconstituted ion chromatogram from GC-MS analysis of the extractable aliphatic hydrocarbon fraction of a diatomaceous ooze from Walvis Ridge (DSDP 75-532-42-3, 173-m sub-bottom depth). Insets *a*, *b* and *c* are expansions of the scan regions 1,050–1,300, 1,900–2,000 and 2,100–2,200, respectively. This sample contains a number of alkyl thiophenes distinguished by the prominence of m/z = 97, 98, 111 or 125 in their mass spectra. Compounds I and II are discussed in the text; their mass spectra for molecular ion fragment mass (M^+) 308 are dominated by m/z of 111 and 98, respectively, and are published in ref. 10. A, D, E and F appear to be C_{18} , C_{19} , C_{25} and C_{27} *n*-alkyl substituted thiophenes. The mass spectrum of F (M^+ = 406) is given in ref. 9. The minor constituents B and C are probably 2,3-dimethyl-5-(2,6,10-trimethylundecyl)-thiophene¹³ and 3,5-dimethyl-2-(3,7,11-trimethyldodecyl)-thiophene, respectively (Fig. 2). Peaks corresponding to *n*-alkanes are designated by their carbon numbers; among other significant components note the abundance of three phytene isomers between scans 700 and 800. The GC-MS conditions were similar to those previously reported²⁹.



thiophenes in shallow, immature sediments (Table 1) suggests, however, that elevated temperatures are not required for their formation. Rather, their occurrence in oceanic sediments is more closely analogous to that of various low-molecular-weight organo-sulphur compounds, including dimethylsulphide and volatile thiophenes²⁵, which are deemed to be direct metabolic products. A biosynthetic origin of higher-molecular-weight thiophenes, such as I, cannot be fully excluded.

A useful method for assessing the origins and extent of thermal maturation of acyclic isoprenoids is the evaluation of the steric configuration at their chiral centres. Comparative GC on a

diethylene glycol succinate/polyethylene glycol succinate (DEGS/PEGs)-coated column²⁶ showed that the naturally occurring compound in a calcareous clay from the Cariaco Trench was composed of a maximum of two of the four possible stereoisomers (Ia–d, Fig. 4) found in the all-isomer synthetic product. It also seems probable from this analysis that the naturally occurring I in the Cariaco Trench sediment is composed, at least in part, of the isomer (3R, 7R; Ia in Fig. 4) expected to derive from isoprenoid biosynthesis, although proof of this requires stereospecific synthesis in the laboratory. The structure and limited stereochemistry of I is consistent with sulphur incorporation into phyta-1,3-diene (IV in Fig. 2) or phytol.

By analogy with the mass spectrum of I (ref. 9), compound II (Table 1, Fig. 1) would correspond to 3-(4,8,12-trimethyltridecyl)-thiophene, which might arise from sulphur incorporation into phyta-1,3(17)-diene (V in Fig. 2) or phytol. Such phytadienes, and hence isoprenoid thiophenes, cannot be artefacts generated during sample preparation because wet extraction methods were used²⁷. Given the widespread occurrence of I and II, and other *n*-alkyl and isoprenyl substituted thiophenes (ref. 13 and S.C.B., unpublished data), in oceanic sediments (Table 1), there appears to be a general metabolic or diagenetic process which results in the introduction of sulphur into lipid moieties. The precise mechanism and biological or chemical agents of this process are unclear, but H_2S or polysulphides may be involved. Perhaps it is a side reaction associated with bacterial sulphate reduction, which certainly occurs or has occurred in the sediments found to contain thiophenes.

For the thiophenes discussed here, the sulphur introduced into the acyclic isoprenoids is bonded intramolecularly. Similar reactions operating in an intermolecular fashion would give rise to sulphur-linked polymeric material within kerogens. Such bonding may survive in petroleum and occur in both their aromatic and asphaltene fractions; especially the latter since it is rich in sulphur. Alternatively, thermal breaking of carbon-sulphur bonds may be a significant contributory process in the generation of petroleum from kerogen. The clear indication that sulphur incorporation can affect organic compounds during early diagenesis represents a major advance in the understanding of interactions at the molecular level between the sulphur and

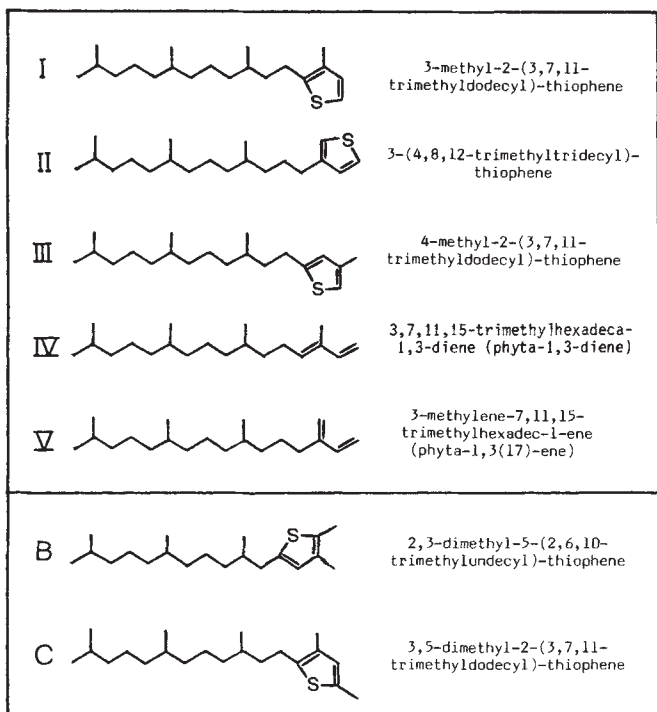


Fig. 2 Structures of compounds cited in the text and figures.

Fig. 3 Summary scheme for synthesis of all-isomer 3-methyl-2-(3,7,11-trimethyldodecyl)-thiophene (I) together with their 4-methyl isomers (III). The methyl substitution at C-3 and C-4 can be distinguished by mass spectrometry, since the compounds show base peaks at $m/z = 111$ and 112 , respectively. Proton-NMR (CDCl_3 , 200 MHz) proved the presence of the different methyl isomers: I gave $\delta_{\text{H}5} = 6.984$ p.p.m., $J = 5.1$ Hz and $\delta_{\text{H}4} = 6.766$ p.p.m., $J = 5.1$ Hz and III gave $\delta_{\text{H}3} = 6.589$ p.p.m. and $\delta_{\text{H}5} = 6.657$ p.p.m. with very small splitting, possibly caused by coupling with the C-4 methyl of the thiophene ring.

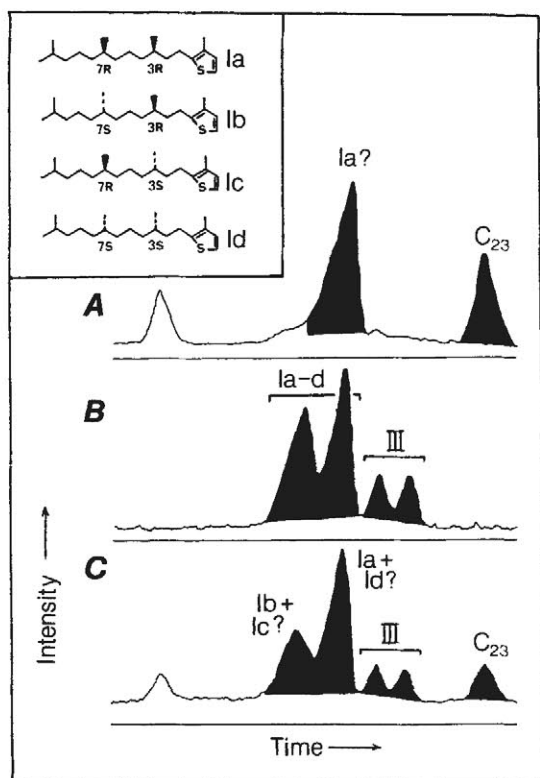
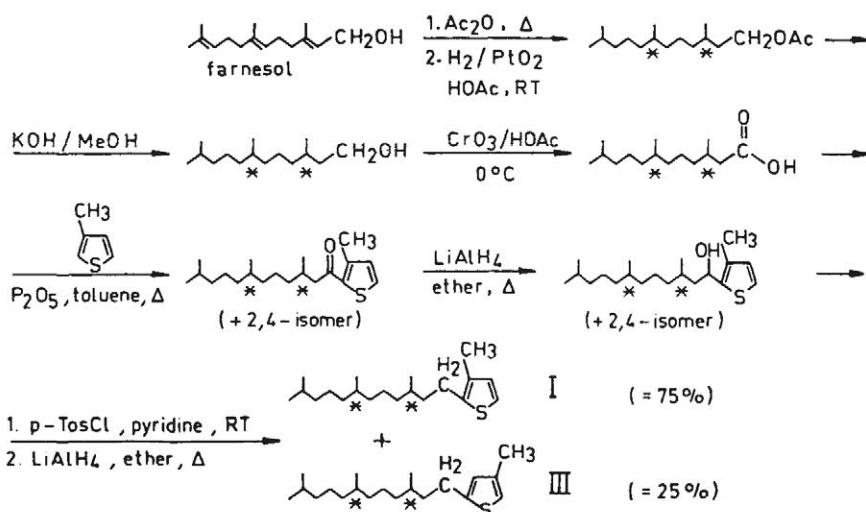


Fig. 4 Partial GC traces using a DEGS/PEGs column of the n -tricosane ($n\text{-C}_{23}$) region of: **A**, aliphatic hydrocarbon fraction of a Pleistocene calcareous clay from the Cariaco Trench (DSDP 15-147C-3-3, 138-m sub-bottom depth). **B**, Products (I and III) of the synthesis shown in Fig. 3. **C**, Coinjection of **A** and **B**. The naturally occurring thiophene enhances the latter peak attributed to I. This evidence can be compared with that from similar GC analyses of acyclic isoprenoids where the isomer with the stereochemistry corresponding to both that of the phytol moiety in chlorophyll- a ³⁰ and archaeobacterial phytanyl ether moieties³¹ occurs in the peak with the greater elution time (for example 7R, 11R in dihydrophytol³¹⁻³⁴ and phytanic acid^{32,34}, and 6R, 10S in phytane³⁵).

Methods. 34-m DEGS/PEGs (3:1) glass capillary column fitted in Carlo Erba FTV2150 chromatograph programmed from 20–109°C at 4°C min⁻¹ and held isothermally at 109°C. Data are acquired and processed using a VG Minichrom data system. Assignment of $n\text{-C}_{23}$ was made by comparison with the elution position of a reference standard. Note that the elution times of the thiophenes relative to n -alkanes are markedly greater with this column than with that (OV-1) employed in the GC-MS analysis (Fig. 1).

carbon cycles within sediments. Also, the occurrence of organo-sulphur compounds indicates that organic matter, like iron²⁸, can act as a sink for sedimentary sulphur. Thus, the identification of acyclic isoprenoid thiophenes in sediments demonstrates a new, significant process in the diagenetic alteration of lipids, and presumably other compound classes.

S.C.B. and C.A.L. are grateful to NERC for GC-MS facilities (grants GR3/2951 and GR3/3758 to Professor G. Eglinton), to the Royal Society for an instrument grant (VG Minichrom system) and to British Petroleum PLC for a studentship (C.A.L.).

Received 2 September 1985; accepted 23 January 1986.

1. Tissot, B. P. & Welte, D. H. *Petroleum Formation and Occurrence* (Springer, Berlin, 1984).
2. Orr, W. L. *Amer. chem. Soc. Div. Fuel Chem. Preprints* **22**, 86–99 (1977).
3. Speers, G. C. & Whitehead, E. V. in *Organic Geochemistry: Methods and Results* (eds Eglinton, G. & Murphy, M. T. J.) 638–675 (Springer, Berlin, 1969).
4. Ho, T. Y., Rogers, M. A., Drushel, H. V. & Koons, C. B. *Bull. Amer. Ass. Petrol. Geol.* **58**, 2338–2348 (1974).
5. Payzant, J. D., Cyr, T. D., Montgomery, D. S. & Strausz, O. P. *Tetrahedron Lett.* **26**, 4175–4178 (1985).
6. Valisollalao, J., Perakis, N., Chappe, B. & Albrecht, P. *Tetrahedron Lett.* **25**, 1183–1186 (1984).
7. Rullkötter, J., von der Dick, H. & Welte, D. H. *Init. Rep. DSDP* **63**, 819–836 (1981).
8. Rullkötter, J., von der Dick, H. & Welte, D. H. *Init. Rep. DSDP* **64**, 837–853 (1982).
9. Rullkötter, J., Mukhopadhyay, P. K. & Welte, D. H. *Init. Rep. DSDP* **75**, 1069–1087 (1984).
10. Rullkötter, J., Mukhopadhyay, P. K., Schaefer, R. G. & Welte, D. H. *Init. Rep. DSDP* **79**, 775–806 (1984).
11. van Graas, G. thesis, Delft Univ. Technol. (1982).
12. Klok, J. thesis, Delft Univ. Technol. (1984).
13. ten Haven, H. L., de Leeuw, J. W. & Schenck, P. A. *Geochim. cosmochim. Acta* **49**, 2181–2191 (1985).
14. Dean, R. A. & Whitehead, E. V. *Tetrahedron Lett.* 768–770 (1961).
15. Bendoraitis, J. G., Brown, B. L. & Hepner, L. S. *Analyt. Chem.* **34**, 49–53 (1962).
16. Brassell, S. C., Wardroper, A. M. K., Thomson, I. D., Maxwell, J. R. & Eglinton, G. *Nature* **290**, 693–696 (1981).
17. Chappe, B., Michaelis, W. & Albrecht, P. in *Advances in Organic Geochemistry 1979* (eds Douglas, A. G. & Maxwell, J. R.) 265–274 (Pergamon, Oxford, 1980).
18. Risatti, J. B., Rowland, S. J., Yon, D. A. & Maxwell, J. R. *Org. Geochem.* **6**, 93–104 (1984).
19. Goossens, H., de Leeuw, J. W., Schenck, P. A. & Brassell, S. C. *Nature* **312**, 440–442 (1984).
20. Kinney, I. W. & Cook, G. L. *Analyt. Chem.* **24**, 1391–1396 (1952).
21. Pomonis, J. G., Fatland, C. L. & Zaylskie, R. G. *J. chem. Engng Data* **21**, 380–385 (1976).
22. Brassell, S. C. & Eglinton, G. in *Advances in Organic Geochemistry 1981* (eds Björjör, M. et al.) 684–697 (Wiley, Chichester, 1983).
23. Brassell, S. C. et al. in *Advances in Organic Geochemistry 1979* (eds Douglas, A. G. & Maxwell, J. R.) 375–392 (Pergamon, Oxford, 1980).
24. Mango, F. D. *Geochim. cosmochim. Acta* **47**, 1433–1441 (1983).
25. Whelan, J. K., Hunt, J. M. & Berman, J. R. *Geochim. cosmochim. Acta* **44**, 1767–1785 (1980).
26. McKenzie, A. S., Patience, R. L., Yon, D. A. & Maxwell, J. R. *Geochim. cosmochim. Acta* **46**, 783–792 (1982).
27. De Leeuw, J. W. et al. in *Advances in Organic Geochemistry 1975* (eds Campos, R. & Goni, J.) 61–79 (Enadisma, Madrid 1977).
28. Berner, R. A. *Geochim. cosmochim. Acta* **48**, 605–615 (1984).
29. Brassell, S. C., Goward, A. P. & Eglinton, G. in *Advances in Organic Geochemistry 1979* (eds Douglas, A. G. & Maxwell, J. R.) 421–426 (Pergamon, Oxford, 1980).
30. Burrell, J. W. K., Garwood, R. F., Jackman, L. M., Oskay, E. & Weedon, B. C. L. *J. chem. Soc. C*, 2144–2154 (1966).
31. Anderson, R., Kates, M., Baedecker, M. J., Kaplan, I. R. & Ackman, R. G. *Geochim. cosmochim. Acta* **41**, 1381–1390 (1977).
32. Brooks, P. W. & Maxwell, J. R. in *Advances in Organic Geochemistry 1973* (eds Tissot, B. P. & Biener, F.) 977–991 (Technip, Paris, 1974).
33. Van Vleet, E. S. & Quinn, J. G. *Geochim. cosmochim. Acta* **43**, 289–303 (1979).
34. Prahl, F. G., Eglinton, G., Corner, E. D. S. & O'Hara, S. C. M. *Science* **224**, 1235–1237 (1984).
35. Patience, R. L., Yon, D. A., Ryback, G. & Maxwell, J. R. in *Advances in Organic Geochemistry 1979* (eds Douglas, A. G. & Maxwell, J. R.) 287–293 (Pergamon, Oxford, 1980).

Floral evidence for Cretaceous chloranthoid angiosperms

E. M. Friis*, P. R. Crane† & K. R. Pedersen*

* Department of Geology, University of Aarhus, DK-8000 Aarhus C, Denmark

† Department of Geology, Field Museum of Natural History, Chicago, Illinois 60605, USA

The hypothesis that multiparted magnolialean flowers retain the largest number of primitive floral characters among living angiosperms has received support from almost 80 years of comparative studies with extant plants¹⁻³. Accumulating fossil data suggest, however, that large *Magnolia*-like flowers were probably preceded in the fossil record by smaller and simpler floral types, some of them perhaps related to the Chloranthaceae, a family generally regarded as being advanced within the Magnoliidae²⁻⁴. The presence of chloranthoid plants early in the history of the angiosperms has been suggested by studies of dispersed pollen⁴⁻⁷ and fossil leaves⁸, but the information on floral structure crucial to assessing the biology of these plants and their relationships within the Chloranthaceae, has so far been unavailable. We now provide new evidence of Cretaceous chloranthoid angiosperms based on fossil androecia, with pollen *in situ*, from the Lower Cretaceous of eastern North America and the Upper Cretaceous of southern Sweden. The Cretaceous material clarifies the homologies of chloranthoid androecial structures and provides an improved basis for interpreting the pollination biology in this enigmatic group of early angiosperms.

The earliest well-documented angiosperm fossils include pollen grains of *Clavatipollenites hughesii*, originally described from the late Barremian and Aptian of southern England^{5,9,10}. Couper⁵ noted the resemblance of these pollen grains to those of the extant chloranthaceous genus *Ascarina*, and detailed similarities in pollen wall structure have subsequently been demonstrated by transmission and scanning electron microscopy of fossil and Recent material⁴. Two other Lower Cretaceous angiosperm pollen types, *Stephanocolpites fredericksburgensis* and *Asteropollis asteroides*, are also closely similar to pollen of extant Chloranthaceae and have been compared to the genera *Chloranthus* and *Hedyosmum*, respectively⁴. Certain Lower Cretaceous leaves also exhibit strong epidermal and architectural similarities to extant Chloranthaceae⁸, although these are not sufficiently well marked to permit an assignment to the modern family.

The material described here provides the first direct evidence of floral structure in early chloranthoid angiosperms. It com-

prises two kinds of androecial structure obtained by sieving from the Lower Cretaceous Patapsco Formation (Potomac Group) of eastern North America and from the Upper Cretaceous of Scania, southern Sweden. The Lower Cretaceous material was recovered from a clay lens at the West Brothers locality in Maryland, dated on palynological evidence as probably early late Albian (Pollen Subzone II-B)¹¹. The associated flora includes wood fragments, logs and numerous angiosperm leaf fossils^{8,12} as well as conifer shoots, cones and seeds, and a variety of angiosperm fruits, seeds, anthers and flowers. The Upper Cretaceous material was recovered from clays and sands in the Höganäs AB kaolin quarry near Åsen, Scania, dated on palynological evidence as of late Santonian or early Campanian age. A diverse assemblage of angiosperm flowers, fruits and seeds has already been reported from the Åsen locality^{13,14}.

The androecial structures from both localities are three-dimensionally preserved as charcoal, and scanning electron microscopy reveals both morphological details and *in situ* pollen. Both androecia are three-lobed with the lateral lobes slightly shorter than the central lobe. In the Lower Cretaceous material each lobe consists of a fully developed stamen with two pairs of small, elongated pollen sacs borne laterally on a fleshy and almost cylindrical connective that is slightly expanded at the apex. The three stamens are fused only at the base (Fig. 1a). The central lobe is ~0.7 mm long, and the total width of the androecium is ~0.5 mm. The anthers contain small, coarsely reticulate pollen grains, ~10 µm long. The pollen apparently has three (or four) colpi, but the precise number of apertures is obscured by folds in the grains and by a thick coating resembling pollenkitt (Fig. 1b). In the Upper Cretaceous fossils only the central lobe of the androecium has four pollen sacs, and the two lateral lobes each bear a single pair of pollen sacs on the adaxial surface (Fig. 1c). The connectives of all three lobes are broad and fleshy, with a short apical expansion. They are fused immediately below the pollen sacs and form a broad laminar structure, ~1-2 mm long and 1-2 mm wide. Pollen grains observed *in situ* are small (~13 µm in diameter), reticulate and spiraperturate (Fig. 1d).

Extant Chloranthaceae comprise approximately 70 species distributed among four genera. *Ascarina* (including *Ascarinopsis*) and *Hedyosmum* have unisexual flowers and an arborescent or shrubby habit, while *Chloranthus* and *Sarcandra* have bisexual flowers and a suffrutescent or herbaceous habit. All genera produce reticulate pollen grains, but within the family there is considerable variation in the number and arrangement of apertures^{4,15,16}. The three-lobed arrangement of both the Lower and Upper Cretaceous androecia, combined with the expanded connectives, small pollen sacs and reticulate pollen, constitutes a unique combination of characters that occurs today

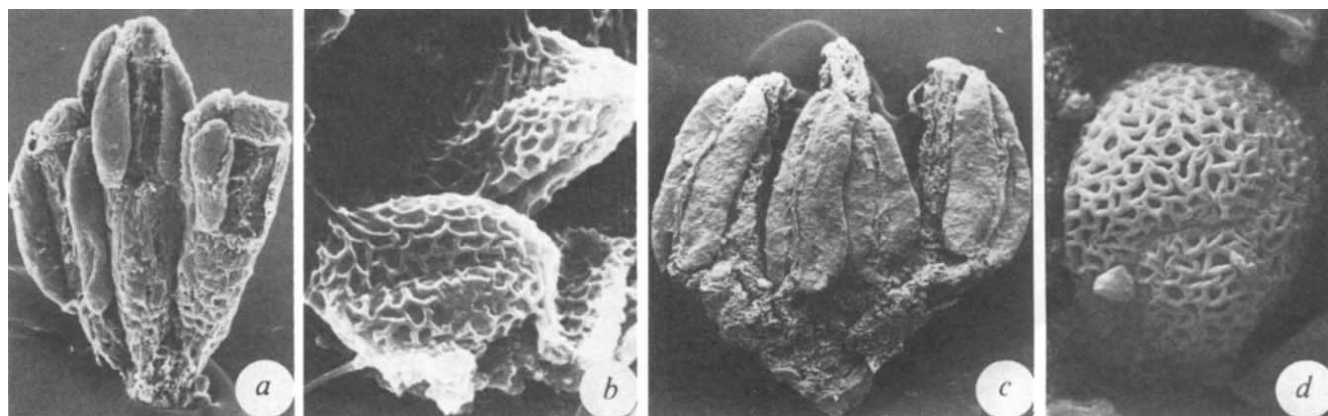


Fig. 1 Scanning electron micrographs of fossil chloranthoid androecia and pollen from the Cretaceous. a, Chloranthoid androecium from the Lower Cretaceous (Upper Albian of eastern North America) ($\times 80$). b, Pollen grains from the androecium shown in a, adhering to the anther by means of a substance resembling pollenkitt ($\times 3,300$). c, Chloranthoid androecium from the Upper Cretaceous (Upper Santonian/Lower Campanian) of southern Sweden ($\times 50$). d, Pollen grains from the androecium shown in c; note the spiral aperture ($\times 3,300$).

only in the genus *Chloranthus*. The androecia of modern *Chloranthus* are more or less laminar and identical to that of the Upper Cretaceous form in the arrangement and number of pollen sacs, with one pair on the lateral lobes and two pairs on the central lobe. Owing to lack of information on the gynoecium, however, the fossils cannot be included with certainty in the extant genus *Chloranthus*. The Lower Cretaceous androecium has a more generalized morphology, and there are no extant species of *Chloranthus* that have lateral lobes with a full complement of four pollen sacs, or that have laterally positioned pollen sacs borne on a more or less cylindrical connective. In *Sarcandra*, however, the androecium consists of a single stamen, closely resembling each lobe of the Lower Cretaceous fossil. The Lower Cretaceous form may thus be related to the ancestral chloranthoids that may have given rise to both *Chloranthus* and *Sarcandra*.

The nature of the *Chloranthus* androecium is the subject of much controversy. Some authors have suggested an origin by lateral fusion of three stamens, one with a fully developed anther consisting of two pairs of pollen sacs, and two with half-anthers each consisting of a single pair of pollen sacs^{3,15}. Other authors have suggested an origin by division of a single stamen that bore four pairs of pollen sacs¹⁵. The structure of the Lower Cretaceous androecium reported here supports the more straightforward view that the *Chloranthus* androecium originated by simple lateral fusion of three normally developed stamens, each of which originally bore four pollen sacs. This notion is also consistent with the available anatomical and ontogenetic data on extant *Chloranthus*¹⁵. The loss of the two pollen sacs in each of the lateral stamens may have occurred with increasing fusion, and a shift to adaxially positioned pollen sacs may have been established as the androecium became more laminar. Both changes may have occurred as part of a trend towards increasingly efficient pollination.

Extant Chloranthaceae include both insect- and wind-pollinated taxa. The bisexual flowers of *Chloranthus* and *Sarcandra* are apparently pollinated by beetles or occasionally flies¹⁷. Insects are attracted to the flowers by the enlarged, fleshy connectives, which become coloured and glossy at anthesis and produce a fruit-like aroma. Pollen production is low and the pollen grains are covered by sticky pollenkit¹⁷. In *Ascarina* and *Hedyosmum* the flowers are unisexual and apparently anemophilous^{17,18}. The pollen grains are dry and produced in large quantities.

The two fossil androecia reported here strongly suggest pollination by insects. In addition to the probable relationship to *Chloranthus* and *Sarcandra*, the presence of fleshy connectives, small pollen grains and a coating of pollenkit all support this interpretation. The number of apertures in the pollen from the Lower Cretaceous androecium is not fully clear, but no tri- or polycolpate pollen of comparable size and sculpture has been recognized among the dispersed pollen described from the Potomac Group. Similarly, spiraperturate pollen comparable to that from the Upper Cretaceous androecia has not been observed in the dispersed pollen flora from Åsen; this also suggests the possibility of low pollen production in our Cretaceous material. In contrast, the relative abundance of *Clavatipollenites* (similar to *Ascarina*) and *Asteropollis* (similar to *Hedyosmum*) in some Lower Cretaceous samples has been used to infer that the plants producing these grains were probably anemophilous⁴. If this interpretation is correct, then chloranthoid angiosperms had already differentiated into anemophilous and entomophilous taxa before the end of the late Cretaceous. This supports the suggestion of an early differentiation of insect- and wind-pollination systems in the angiosperms, based on slightly younger material¹⁹.

The evidence now available on Cretaceous Chloranthaceae demonstrates conclusively the importance of this family in the earliest phases of angiosperm diversification. At least some early angiosperms had small, inconspicuous flowers and evidence

suggests that both wind and insect pollination were established very early in angiosperm evolution.

We thank Professor W. L. Crepet, Professor D. L. Dilcher, Dr P. K. Endress and Dr K. Nixon for helpful discussion. We acknowledge the receipt of a Niels Bohr fellowship (E.M.F.) and support from the NATO Senior Fellowship Scheme (E.M.F.), the Field Museum Department of Geology Visiting Scientist Program (K.R.P.) and NSF grant BSR-8314592 (P.R.C.).

Received 17 October 1985; accepted 27 January 1986.

- Arber, E. A. N. & Parkin, J. J. *Linn. Soc.* **38**, 29–80 (1907).
- Takhtajan, A. *Flowering Plants, Origin and Dispersal*, 1–310 (Oliver & Boyd, Edinburgh, 1969).
- Cronquist, A. *An Integrated System of Classification of Flowering Plants*, 1–1262 (Columbia University Press, 1981).
- Walker, J. W. & Walker, A. G. *Ann. M. bot. Gdn* **71**, 464–521 (1984).
- Couper, R. A. *Palaeontographica B103*, 77–179 (1958).
- Doyle, J. A. *J. Arnold Arbor.* **50**, 1–35 (1969).
- Muller, J. *Bot. Rev.* **47**, 1–142 (1981).
- Upchurch, G. R. *Ann. M. bot. Gdn* **71**, 522–550 (1984).
- Kemp, E. M. *Palaeontology* **11**, 421–434 (1968).
- Hughes, N. R., Drewry, G. E. & Laing, J. F. *Palaeontology* **22**, 513–535 (1979).
- Doyle, J. A. & Hickey, L. J. in *Origin and Early Evolution of Angiosperms* (ed. Beck, C. B.) 139–206 (Columbia University Press, 1976).
- Hickey, L. J. in *Cretaceous and Tertiary Stratigraphy, Paleontology and Structure, South-western Maryland and Northeastern Virginia* (eds Frederiksen, N. O. & Kraft, K.) 193–209 (U.S. Geological Survey, Reston, 1984).
- Friis, E. M. *Ann. M. bot. Gdn* **71**, 403–418 (1984).
- Friis, E. M. *Biol. Skr.* **25**, 1–37 (1985).
- Swamy, B. G. L. *J. Arnold Arbor.* **34**, 375–408 (1953).
- Erdtman, G. *Pollen Morphology and Plant Taxonomy*, 1–539 (Almqvist & Wiksell, Stockholm, 1952).
- Endress, P. K. *Pl. Syst. Evol.* (in the press).
- van der Hammen, T. & Gonzales, E. *Leid. geol. Meded.* **25**, 261–315 (1960).
- Dilcher, D. L. *Rev. Palaeobot. Palynol.* **27**, 291–328 (1979).

Early forest clearance and environmental degradation in south-west Uganda

Alan Hamilton*, David Taylor* & J. C. Vogel†

* Department of Environmental Studies, University of Ulster, Coleraine, Co. Londonderry BT52 1SA, UK

† National Physical Research Laboratory, PO Box 395, Pretoria 001, South Africa

Evidence for early agriculture can be obtained from pollen profiles indicating forest clearance¹. The practice of cultivation is widely believed to have been introduced into the interlacustrine region of central Africa by Bantu-speaking iron-workers², present by 2,000 yr BP (ref. 3) or, disputedly^{2–4}, 2,600 yr BP (ref. 5). There are, however, archaeological and linguistic indications that cultivation may have begun earlier^{3,4,6,7}. Published pollen diagrams which clearly show forest clearance in East Africa are either poorly dated⁸ or place forest destruction¹ after 2,000 yr BP; possible palynological signs of previous clearance^{9,10} are open to other interpretations, notably climatic change⁷. We have now analysed the pollen and charcoal of a peat core from Ahakagyazi Swamp in south-west Uganda using radiocarbon for dating. The results suggest a history of forest clearance stretching back beyond 4,800 yr BP. This is the earliest evidence for cultivation reported from tropical Africa.

Ahakagyazi Swamp (1,830 m; Fig. 1) is surrounded by steep hillslopes, covered today with a patchwork of small fields and short pasture¹¹. Few indigenous woody plants remain, but the rainfall is 1,125 mm yr⁻¹ (ref. 12) and the natural vegetation has been reconstructed as moist lower montane forest¹³, similar to that in early Bwindi Forest¹¹. Erosion of the infertile soils developed from Precambrian argillitic rocks¹² is a major agricultural problem¹⁴. The swamp vegetation is *Syzygium* forest with a marginal zone of sedges or *Typha*¹¹.

The 10-m core, collected with a Hiller borer, is mainly wood peat, with herbaceous macrofossils and clay common in two horizons towards the base and herbaceous fossils also present

Fig. 1 Part of the Rukiga Highlands, south-west Uganda, showing Ahakageyezi and other sites^{8,23} used for pollen analysis. Bwindi is moist lower montane forest, grading into lowland forest towards the north-west; the vegetation varies topographically and altitudinally¹¹. Other forests shown are secondary.

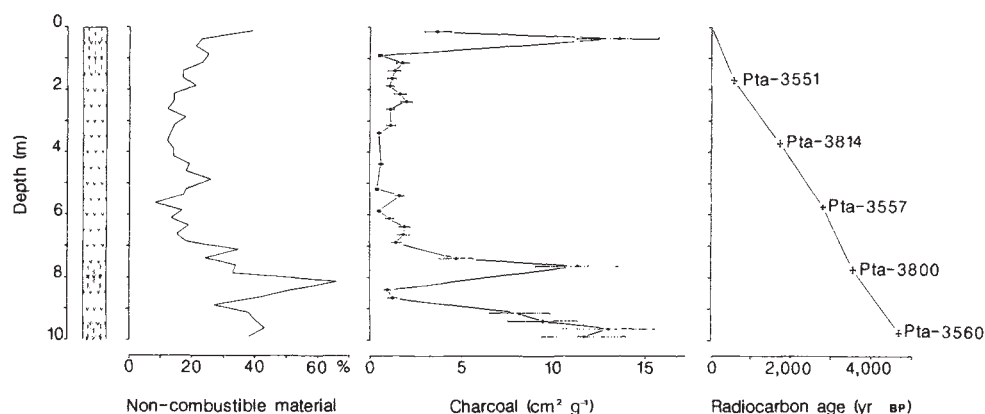
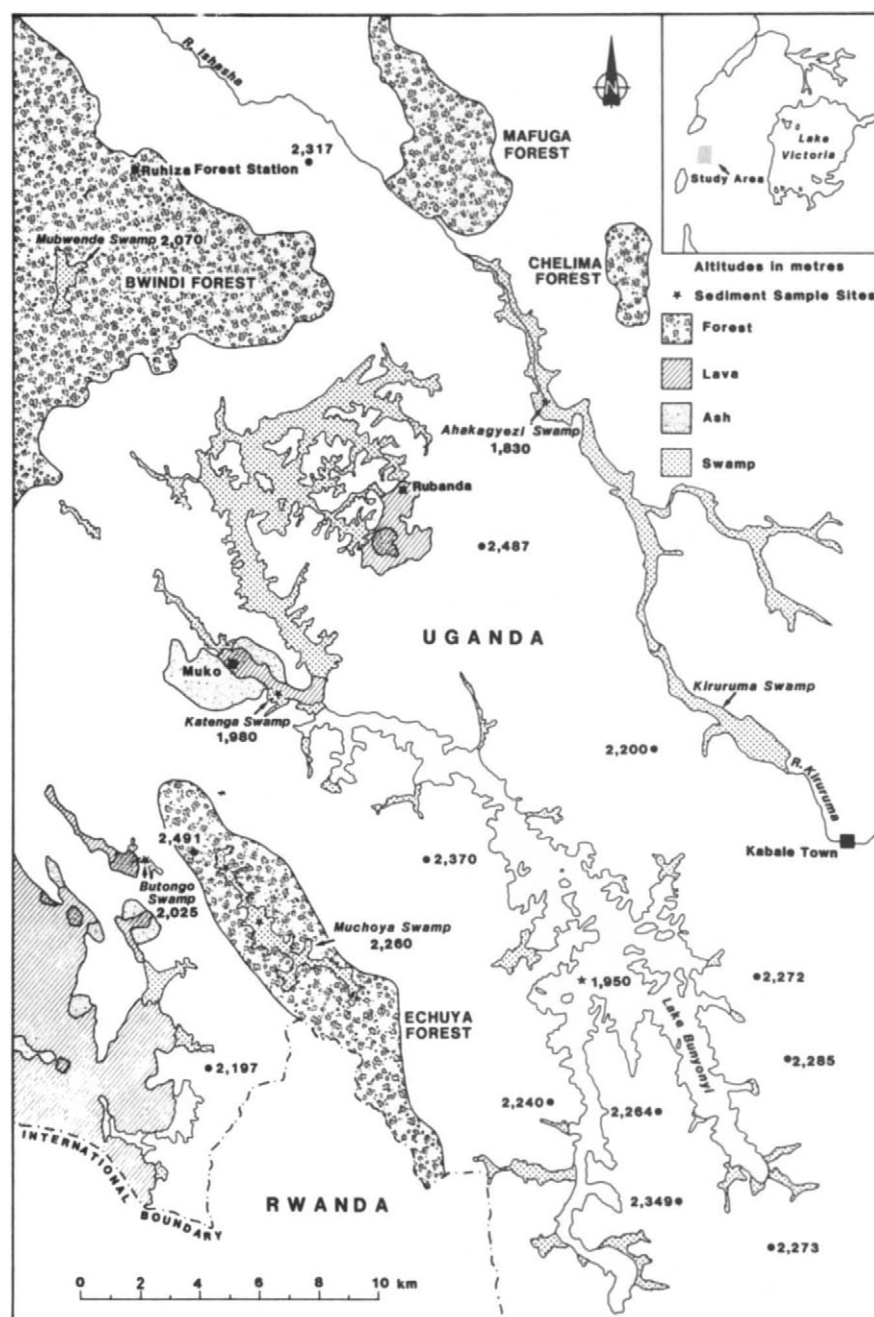


Fig. 2 Stratigraphy and sediment properties of a core from Ahakageyezi Swamp. Non-combustible material determined by ignition of dried samples at 550°C. Charcoal abundance estimated by the point-count method²⁴; bars indicate 95% confidence limits of the counts. Radiocarbon determined on whole peat samples, pretreated with dilute acid; dates calculated using a half life of 5,568 yr, corrected for isotopic fractionation. ▫, Wood peat; ▨, herbaceous peat; ▩, clay.

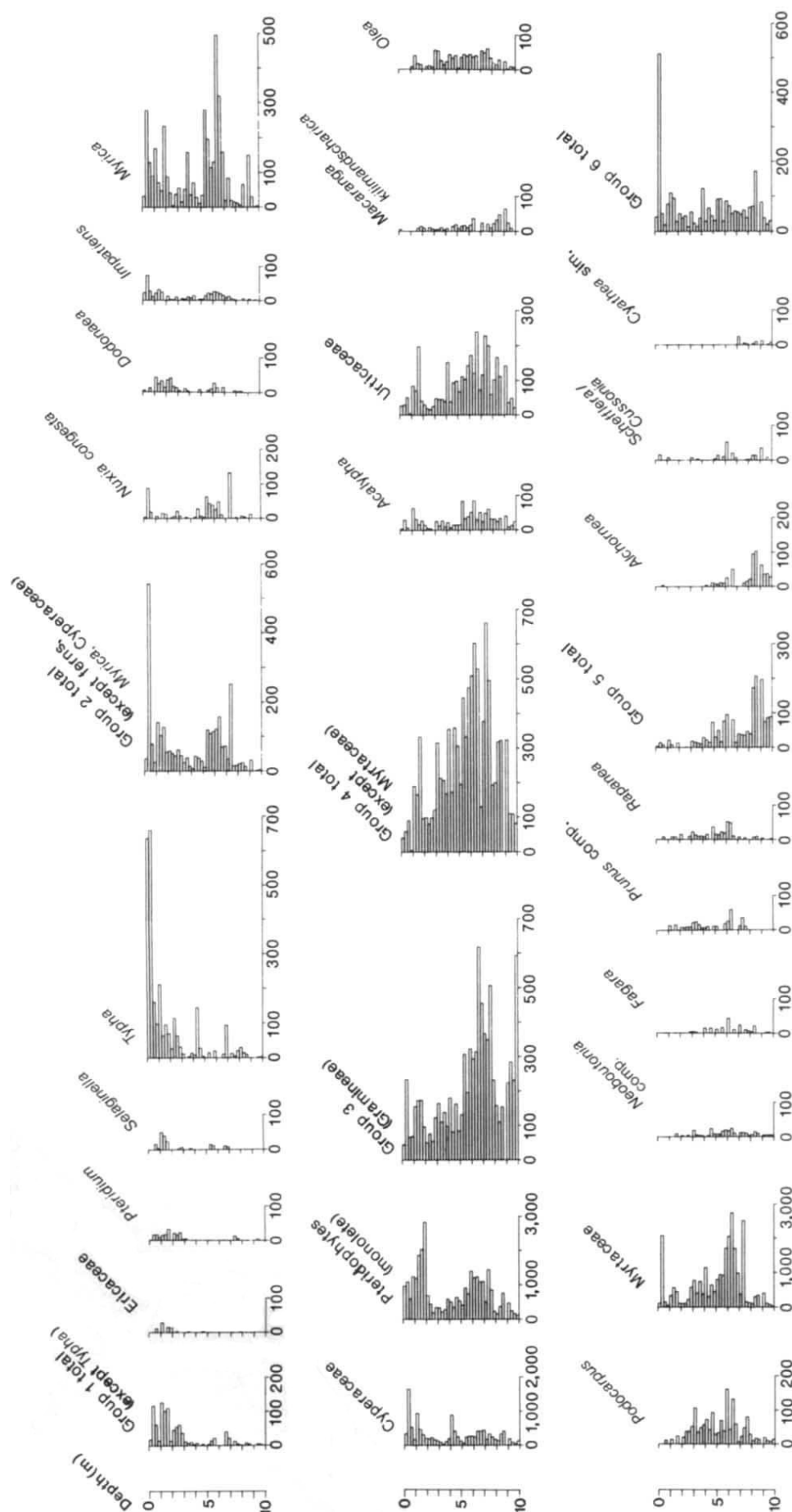


Fig. 3 Ahakagezi: pollen diagram with group totals and selected individual taxa. Pollen abundance expressed as grains deposited per cm² per yr. Level 8.85–8.90 m omitted because of suspected error in estimating pollen concentration. The group totals omit *Typha* (group 1), Cyperaceae, *Myrica*, pteridophytes (monolete) (group 2) and Myrtaceae (group 4), all of which attain high values; if included, they would unduly dominate the group totals. All identified pollen types with an abundance of >0.4% of total pollen at any one level are on the figure or listed below. Additional pollen types: group 1, *Clematis*, Trilete spores (smooth); group 2, *Bidens* sim., Crassulaceae comp., *Polyscias*, *Rumex*, *Stephania*; group 4, *Celtis* sim., *Maesa* comp., *Polypodium* sim., Umbelliferae; group 5, *Anthoceros*, *Anthocleista*, Ebenaceae, Lorantheae, *Musanga/Myrianthus*, *Parinari*; group 6, Acanthaceae, *Achyranthes* comp., *Begonia*, *Faurea* comp., *Helichrys* sim., *Hydrocotyle* sim., *Ilex*, *Laurembergia*, *Olinia*, *Potamogeton*. Ecological interpretation of group 6, characterized by little directional change, is not given.

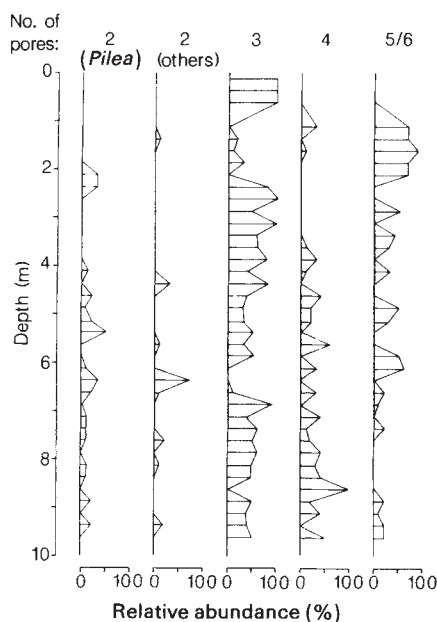


Fig. 4 Urticaceae pollen: percentage abundance of various types in the Ahakagezi core. Those most clearly indicative of forest are *Pilea* and four-porate grains (produced mainly by *Urera hypselodendron*)²⁵.

above 1.5 m (Fig. 2). The horizons containing clay or herbaceous fossils are associated palynologically with periods of forest clearance or intensified agriculture and are believed to signify enhanced soil erosion. Microscopic charcoal is abundant at some levels, especially in and just above the clay-rich strata. Radiocarbon measurements provide excellent dating control (Table 1).

Samples were prepared for pollen analysis by a standard procedure including HF treatment and acetolysis¹⁵; pollen influx rates were estimated using the *Lycopodium* spore method¹⁶. Pollen preservation is generally excellent. Pollen types are grouped here according to patterns of change in accumulation rate, regardless of absolute values (Fig. 3).

Group 1, increasing above ~3 m (1,300 yr BP), includes pollen types derived from plants of agricultural land and scrub^{1,17} (Ericaceae, *Pteridium*, *Vernonia*) and from swamp species associated with heavy siltation¹⁸ (*Selaginella*, *Typha*). This suggests marked disturbance of terrestrial vegetation, leading to soil erosion and consequent vegetational change in the swamp.

Group 2, with two peaks, one similar to that of group 1 and the other at ~5.25–7.5 m (2,500–3,400 yr BP), contains pollen types of plants associated with human disturbance (*Bidens*, *Dodonaea*, *Nuxia*, *Polyscias*, *Rumex*) and pollen of swamp margin plants (Cyperaceae, *Impatiens*, ferns). This indicates a phase of marked human disturbance preceding that already noted.

Gramineae is the only member of group 3, with high values at 5.25–8 m (2,500–3,700 yr BP) and 9–10 m (4,200–4,800 yr BP). Although grasses grow in many habitats^{10,19}, they are rare in the swamp today and these peaks are probably due to spread of secondary grassland replacing forest. Indicators of disturbance in groups 1 and 2 are uncommon at 9–10 m, perhaps because ruderals found establishment difficult under luxuriant grass growth on undegraded forest soils; the low grass values above 5.25 m could be due to declining soil fertility. Increases¹⁰ and decreases¹ of Gramineae pollen with human disturbance in East Africa have been reported, but not previously sequentially at one site.

Group 4 shows an early increase and a later decline. Most parent species are forest trees. Increases in the swamp forest Myrtaceae (*Syzygium*) and the upper slope trees *Olea* and *Podocarpus* suggest a drier climate after 3,500–3,600 yr BP. *Podocarpus* expansion elsewhere in eastern Africa at about this

Table 1 Dating control provided by ¹⁴C measurements

Depth (m)	Age (yr BP)	δ ¹³ C (‰)	Laboratory reference
1.65–1.85	580 ± 50	–27.0	Pta-3551
3.65–3.85	1,690 ± 40	–26.0	Pta-3814
5.70–5.85	2,800 ± 60	–26.9	Pta-3557
7.65–7.85	3,520 ± 50	–18.2	Pta-3800
9.65–9.85	4,670 ± 60	–15.9	Pta-3560

time has been interpreted as due to decreased rainfall¹. Increases in various early successional trees (*Fagara*, *Maesa*, *Neoboutonia*, *Prunus*) could be due to forest disturbance. The falls in abundance of group 4 pollen types towards the surface provide further evidence for forest destruction; lower-slope plants (for example, *Fagara*) tend to decline before upper-slope plants (such as *Podocarpus*), with final clearance of ridge forest at 900 yr BP. Forest Urticaceae pollen becomes rare above 3.5–4 m (1,600–1,800 yr BP) (Fig. 4).

Group 5, with high values at ~8.25–9.25 m (3,800–4,400 yr BP) and a secondary peak at ~4.75–6.75 m (2,200–3,200 yr BP), comprises mainly pollen types of lower-slope forest trees. Little valley forest remained after 2,200 yr BP.

The first two phases of forest reduction were associated with soil erosion and burning. Lightning or volcanism are unlikely causes of such major destruction^{20,21} and in any case volcanism was earlier in this area⁸. Hunting in moist African forests does not involve burning²² and the forest destruction is attributed to the clearance of land for cultivation. Unfortunately, as with other East African pollen diagrams^{1,9}, pollen grains from cultivated plants have not been definitely identified. Charcoal manufacture for iron smelting consumes large quantities of wood⁵ and this could have contributed to the destruction of trees, perhaps including *Syzygium*⁵, during the past two millennia.

This record shows a long history of forest clearance and soil erosion, interrupted by periods of partial forest recovery. Major degradation of the environment by man occurred early at this locality.

We thank the Government of Uganda for permission to undertake this research. We also thank H. Taligoola, A. Katende, D. Baziya, F. Butera, M. Byamugisha, D. Mutigah, A. Mwesigye and C. Sabiti for help in Uganda and W. Magowan, K. McDade, N. McDowell, B. O'Kane, S. Tinkler and C. Todd for assistance at Coleraine. This work was supported by the NERC, the Royal Society and the University of Ulster.

Received 19 June 1985; accepted 6 January 1986.

- Hamilton, A. C. *Environmental History of East Africa* (Academic, London, 1982).
- van der Merwe, M. J. in *The Coming of the Iron Age* (eds Wertim, T. A. & Muhly, J. D.) 463–506 (Yale University Press, New Haven, 1980).
- van Noten, F. *Azania* 14, 61–80 (1979).
- Ehret, C. & Posnansky, M. (eds) *The Archaeological and Linguistic Reconstruction of African History* (University of California Press, Berkeley, 1982).
- Schmidt, P. R. *Tanzania Notes & Records*, 84–85, 77–94 (1980).
- Robertshaw, P., Collett, D., Gifford, D. & Mbae, N. B. *Azania* 18, 1–43 (1983).
- Robertshaw, P. & Collett, D. P. *Wild Archaeol.* 15, 67–78 (1983).
- Morrison, M. E. S. & Hamilton, A. C. *J. Ecol.* 62, 1–31 (1974).
- Kendall, R. L. *Ecol. Monogr.* 39, 121–176 (1969).
- Livingstone, D. A. *Ecol. Monogr.* 37, 25–52 (1967).
- Hamilton, A. C. *Uganda J.* 33, 175–199 (1969).
- Atlas of Uganda* 2nd edn (Department of Lands and Surveys, Entebbe, 1967).
- Langdale-Brown, I., Osmaston, H. A. & Wilson, J. G. in *The Vegetation of Uganda and its Bearing on Land-Use* (Government Printer, Uganda, 1964).
- Kururagire, A. R. *Uganda J.* 33, 59–64 (1969).
- Faegri, K. & Iversen, J. *Textbook of Pollen Analysis* 2nd edn (Munksgaard, Copenhagen, 1964).
- Stockmarr, J. *Pollen Spores* 13, 615–621 (1971).
- Hamilton, A. C. *Palaeoecol. Afr.* 7, 45–149 (1972).
- Deuse, P. *Inst. natn. tech. Scient. Publ.* 4 (I.N.R.S., Butare, Rwanda, 1966).
- Coetzee, J. A. *Palaeoecol. Afr.* 3, 1–146 (1967).
- Dubois, C. G. B. *Uganda J.* 23, 118–123 (1959).
- Rose Innes, R. *Proc. 11th Tall Timbers Fire Ecology Conf.*, 147–173 (Tall Timbers Research Station, Tallahassee, Florida, 1972).
- Turnbull, C. M. *The Forest People* (Cape, London, 1961).
- Morrison, M. E. S. *J. Ecol.* 56, 363–384 (1968).
- Clark, R. L. *Pollen Spores* 24, 523–535 (1982).
- Hamilton, A. C. *Pollen Spores* 18, 27–66 (1976).

A lethal mutation in mice eliminates the slow calcium current in skeletal muscle cells

Kurt G. Beam, C. Michael Knudson
& Jeanne A. Powell*

Department of Physiology and Biophysics, University of Iowa,
Iowa City, Iowa 52242, USA

* Department of Biological Sciences, Smith College, Northampton,
Massachusetts 01063, USA

Contraction of a vertebrate skeletal muscle fibre is triggered by electrical depolarization of sarcolemmal infoldings termed transverse-tubules (t-tubules), which in turn causes the release of calcium from an internal store, the sarcoplasmic reticulum (SR)^{1,2}. The mechanism that links t-tubular depolarization to SR calcium release remains poorly understood. In principle, this link might be provided by the prominent slow calcium current that has been described in skeletal muscle cells of adult frogs^{3,4} and rats⁵. However, blocking this current does not abolish the depolarization-induced contractile responses of frog muscle⁶, and the function of this slow calcium current is unknown. Here we describe measurements of calcium currents in developing skeletal muscle cells of normal rats and mice, and of mice with muscular dysgenesis, a mutation⁷ that causes excitation-contraction (E-C) coupling to fail⁸. We find that a slow calcium current is present in skeletal muscle cells of normal animals but absent from skeletal muscle cells of mutant animals. The effect of the mutation is specific to the slow calcium current of skeletal muscle; a fast calcium current is present in developing skeletal muscle cells of both normal and mutant animals, and slow calcium currents are present in cardiac and sensory neurones of mutant animals. We believe this to be the first report of a mutation affecting calcium currents in a multicellular organism. The effects of the mutation raise important questions about the relationship between the slow calcium current and skeletal muscle E-C coupling.

Calcium currents were recorded using the whole-cell patch-clamp technique⁹ in skeletal muscle, cardiac and neuronal cells obtained from fetal and newborn rats and mice. Figure 1a illustrates the presence of two kinetically distinct inward currents in a rat myotube grown in primary tissue culture. Test potentials in the range -40 to -10 mV elicited a fast, transient, inward current (I_{fast}). Test pulses to potentials of ≥ 0 mV activated an additional, slower, maintained, inward current (I_{slow}), which resembles the slow calcium current of adult skeletal muscle. Both I_{fast} and I_{slow} are calcium currents as they are: (1) observed in Na-free solutions containing 10 μM tetrodotoxin (TTX), (2) abolished by removing calcium from the bathing medium, and (3) blocked by cadmium (see below). I_{fast} and I_{slow} are seen not only in myotubes in primary tissue culture, but also in enzymatically dissociated fibres of muscles that have been acutely removed from neonatal rats (Fig. 1b). The two calcium currents of developing muscle are kinetically similar to the fast and slow calcium currents previously described in cardiac muscle^{10,11}, smooth muscle¹², neuronal cell lines^{13,14} and sensory neurones¹⁵⁻¹⁸. In most of the myotubes we examined, both I_{fast} and I_{slow} were present, although their magnitudes varied considerably between cells. In 20 rat and mouse myotubes examined in 10 mM external calcium, the peak I_{fast} averaged 0.7 ± 0.6 pA pF⁻¹ (mean $\pm$ s.d.) and the peak I_{slow} 8.2 ± 4.9 pA pF⁻¹ (currents normalized by linear cell capacitance). For 13 myotubes in 50 mM external calcium, peak I_{fast} and I_{slow} were 2.1 ± 1.6 and 6.6 ± 4.4 pA pF⁻¹, respectively. In 7 of these 33 cells, I_{fast} was too small to be detected, whereas in only 2 of them was I_{slow} undetectable.

I_{fast} and I_{slow} can be distinguished not only by kinetics and voltage-dependence of activation, but also by the effects of

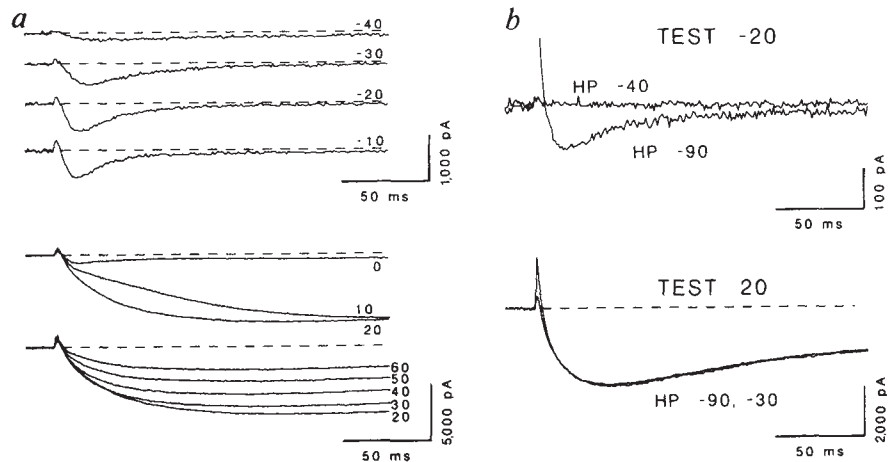
holding potential. Figure 1b illustrates, for an enzymatically dissociated fibre, that changing the holding potential from -90 to -40 mV inactivates the transient I_{fast} elicited by a test pulse to -20 mV; a similar change in holding potential (from -90 to -30 mV) has essentially no effect on I_{slow} elicited by a test pulse to +30 mV. The effects of holding potential on I_{fast} and I_{slow} in myotubes were similar to those illustrated in Fig. 1b.

Calcium channel antagonists provide another means of distinguishing between I_{fast} and I_{slow} . High concentrations (≥ 1 mM) of cadmium block both I_{fast} and I_{slow} , but lower concentrations (≤ 100 μM) produce a significantly greater block of I_{slow} than of I_{fast} . Similarly, 5 μM nitrendipine (a dihydropyridine Ca-channel antagonist) causes a $\sim 75\%$ reduction in I_{slow} , with little effect on I_{fast} . A greater cadmium and dihydropyridine sensitivity has also been described for the slow calcium currents in cardiac cells^{10,11}, a neuronal cell line¹⁴ and sensory neurones^{17,18}. Thus, the voltage-dependence, kinetics and antagonist sensitivity of the fast and slow calcium currents in skeletal muscle all resemble those of fast and slow calcium currents in heart and nerve.

Muscular dysgenesis is a single, recessive, lethal, autosomal mutation (*mdg*) in mice⁷. In dysgenic muscle the sarcolemma can still generate action potentials⁸, and the SR retains the ability to sequester calcium and release it in response to caffeine¹⁹, but sarcolemmal depolarization fails to elicit SR calcium release²⁰. Figure 2 demonstrates that this genetic defect of E-C coupling is associated with the specific absence of I_{slow} . Figure 2a shows calcium currents obtained from a dysgenic myotube. As in normal muscle cells, relatively weak depolarizations elicit I_{fast} ; however, stronger depolarizations fail to elicit I_{slow} , so that even for these strong depolarizations only I_{fast} is present. Altogether, we examined calcium currents in 15 dysgenic myotubes; in none of these was a measurable I_{slow} present. In these same cells the peak I_{fast} was 1.0 ± 0.5 ($n=8$) in 10 mM external calcium and 1.8 ± 1.3 pA pF⁻¹ ($n=7$) in 50 mM Ca, values similar to those in normal myotubes (see above). Figure 2b illustrates the effect of changing the holding potential on the currents elicited in a dysgenic myotube by test pulses to either -10 or +30 mV. When the holding potential was -80 mV, both test pulses elicited I_{fast} (prominent at -10 mV; partially obscured by noise at +30 mV). The same test pulses failed to elicit I_{fast} when the holding potential was -30 mV. Thus, the inactivation of I_{fast} with holding potential is similar in dysgenic myotubes to that in normal skeletal muscle cells. Importantly, the steady-state current for the +30-mV test pulse is in quite good agreement for the -80 and -30 mV holding potentials, which argues against the possibility that large nonlinearities in leak current disguise the presence of I_{slow} in dysgenic myotubes. Thus, the *mdg* mutation seems specific in its effect: I_{slow} is absent but I_{fast} is present with normal properties. Additionally, dysgenic myotubes possess sodium currents of normal appearance (data not shown).

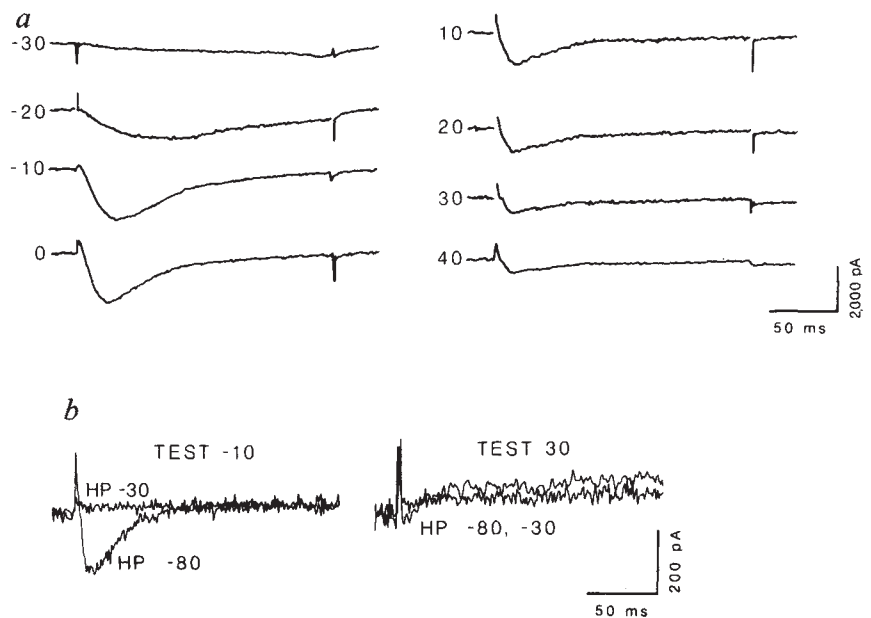
The many similarities between calcium currents in skeletal muscle and those in sensory neurones and cardiac cells prompted us to determine whether the *mdg* mutation also eliminates slow calcium current in these tissues. In control experiments on heart and nerve from non-mutant animals, we observed two clearly distinct components of current resembling I_{fast} and I_{slow} of developing skeletal muscle. (A third calcium current, termed N, has been described for sensory neurones¹⁸, but in our experimental conditions this current was not prominent.) Figure 3 illustrates calcium currents recorded in cardiac myocytes and dorsal root ganglion neurones from mutant animals. In the dysgenic cardiac myocyte (Fig. 3a), a 0-mV test pulse applied from a -80-mV holding potential activated both I_{fast} and I_{slow} ; when the same test pulse was used but the holding potential was changed to -40 mV, I_{fast} was inactivated but I_{slow} was little affected. This change in holding potential also had little effect on the much larger I_{slow} that was activated by a test pulse to +30 mV. Similarly, in the dysgenic neurone (Fig. 3b), a test pulse to -20 mV elicited a predominantly transient (I_{fast})

Fig. 1 Developing skeletal muscle cells possess two distinct calcium currents. *a*, Transient (I_{fast}) and maintained (I_{slow}) calcium currents in a rat myotube. The cell was held at a potential of -90 mV from which it was depolarized to the test potentials indicated to the right of each trace. Test depolarizations of -40 to -10 mV selectively activate I_{fast} ; test potentials of 0 mV and above activate an additional current, I_{slow} . Owing to its much larger size, I_{slow} masks the presence of I_{fast} for test potentials ≥ 20 mV. Rat myotube (A07), 12 days in culture, 10 mM Ca + 10 μ M TTX, holding potential (HP) -90 mV, linear capacitance (C) = 340 pF. *b*, Depolarized holding potentials selectively inactivate I_{fast} . Single fibre (D18), enzymatically dissociated²⁸ from the flexor digitorum brevis muscle of a 13-day rat, 10 mM Ca + 10 μ M TTX, C = 290 pF. Test potentials of -20 and $+20$ mV were applied.



Methods. Currents were recorded with the whole-cell variant of the patch-clamp technique⁹. Unless otherwise indicated, illustrated test currents have been corrected digitally for linear components of capacitive and leak current. The pipette contained (in mM): 140 Cs-aspartate, 5 Mg-aspartate, 10 Cs₂EGTA, 10 HEPES (adjusted to pH 7.4 with CsOH). External solutions, designated by their divalent cation, contained (in mM) 10 Ca: 145 tetraethylammonium ion (TEA), 10 Ca; 50 Ca: 85 TEA, 50 Ca; isotonic Ba: 105 Ba. All external solutions contained 10 mM HEPES (pH 7.4 with CsOH) and Br and/or Cl as anions. Experiments were done at room temperature (20°C). Primary cultures of myoblasts were prepared from fore- and hindlimbs of late-term fetal or newborn rats and mice. The limb muscles were minced finely and incubated at 37°C for 40 min in physiological saline⁵ containing collagenase (2 mg ml⁻¹, Sigma Type I). After filtration and centrifugation to remove large debris, the cell suspension was pre-plated for 1 h in glass to remove rapidly adhering cells and then plated onto 35 -mm Falcon 'primaria' dishes in 'plating medium' containing (v/v) 80% Dulbecco's modified Eagle's medium with 4.5 g l⁻¹ glucose (DMEM), 10% horse serum, 10% calf serum. After 4 days the plating medium was replaced with 'maintenance medium' (90% DMEM, 10% horse serum). All culture media contained penicillin ($1,000$ U ml⁻¹) and streptomycin (1 mg ml⁻¹). Cardiac myocytes were prepared from hearts of newborn mice essentially as described in ref. 29, suspended in plating medium supplemented with 0.1 mM 8-bromo-cyclic AMP, and plated onto polylysine-coated culture dishes. Dorsal root ganglia were removed from either 14 -day-old fetuses or newborn animals, and incubated for 20 min at 37°C in physiological saline containing 0.25% trypsin (Sigma type IIIS). The ganglia were washed, disaggregated, suspended in plating medium enriched with 50 ng ml⁻¹ 7S nerve growth factor (Collaborative Research), and plated onto polylysine-coated culture dishes. Cultures were maintained at 37°C in a 95% air- 5% CO₂, water-saturated atmosphere.

Fig. 2 Skeletal muscle cells of mutant animals specifically lack I_{slow} . *a*, Presence of I_{fast} and absence of I_{slow} in dysgenic myotube. Cell MD158, 6 days in culture, 50 mM Ca, HP = -90 mV, C = 675 pF. *b*, Inactivation of I_{fast} with holding potential in a dysgenic myotube. Cell E11, 10 days in culture, 10 mM Ca + 10 μ M TTX, C = 100 pF.



inward current that was inactivated at a holding potential of -40 mV, whereas a test pulse to $+20$ mV elicited a predominantly maintained (I_{slow}) inward current that was similar for holding potentials of -80 and -40 mV. Thus, both an I_{fast} and an I_{slow} component of current are present in cardiac and neuronal cells of mutant animals, and the effect of the *mdg* mutation seems to be specific for the slow calcium current of skeletal muscle.

The results of recent biochemical experiments²¹ agree with our electrophysiological measurements in demonstrating that the *mdg* mutation affects skeletal muscle but not cardiac cells. Thus, in cardiac tissue the number of binding sites for PN 200-110 (a dihydropyridine calcium-channel blocker) is the

same in normal and dysgenic animals. By contrast, in skeletal muscle of mutant animals, the number of PN 200-110 binding sites is reduced about fivefold relative to the control. The identity of the substantial number of dihydropyridine binding sites that remain in dysgenic skeletal muscle is uncertain. If they represent functional slow calcium channels, the dysgenic myotubes would have an I_{slow} about one-fifth the size of I_{slow} in normal cells, which should have been detectable by our electrophysiological measurements. Our failure to detect I_{slow} suggests that the dihydropyridine binding sites in dysgenic skeletal muscle represent non-functional slow calcium channels or an unrelated component of the muscle membrane.

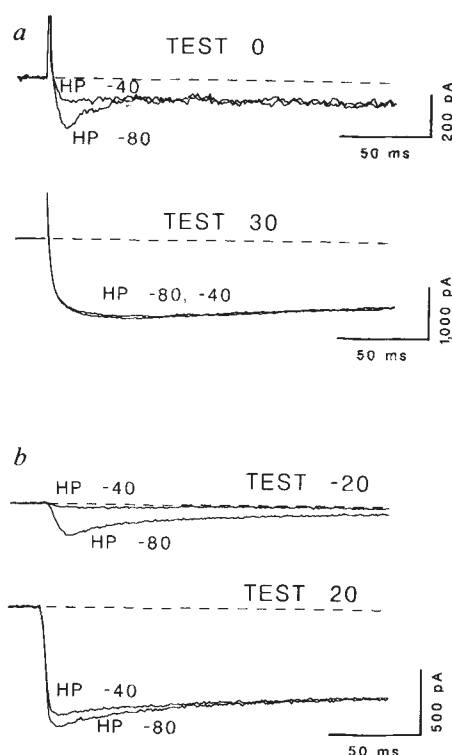


Fig. 3 Cardiac and neuronal cells of mutant animals possess both I_{fast} and I_{slow} . *a*, Dysgenic cardiac myocyte (B59), 8 days in culture, isotonic Ba + 10 μ M TTX, C = 46 pF, no leak subtraction. *b*, Dorsal root ganglion neurone (C44) from a 15-day dysgenic fetus, 3 days in culture, 10 mM Ca + 10 μ M TTX, C = 40 pF.

Why does the *mdg* mutation result in the absence of slow calcium current in skeletal muscle? Among the possibilities are that the mutation directly alters the structural gene for the slow calcium channel, that it interferes with transcription of the gene, that it disrupts the post-translational processing and/or insertion of the channel protein, or that it affects a soluble regulatory protein necessary for channel function. The last of these has been shown to be the case for the paramyotonic mutant *cnrC*, which lacks the calcium current present in wild-type animals²². The presence of slow calcium currents in dysgenic cardiac and neuronal cells suggests the possibility that slow calcium channels in these tissues are encoded by different structural genes from those in skeletal muscle, even though the currents are similar in the three tissues. Alternatively, if the same gene codes for the slow calcium channel in all three tissues, the expression of the slow current in skeletal muscle must involve tissue-specific processes that do not occur in cardiac or neuronal cells.

A second major question raised by our results is the relationship between the absence of I_{slow} and the failure of E-C coupling. One explanation would be to suppose that calcium entry via I_{slow} is a necessary trigger for SR calcium release, but this seems unlikely considering the current's slow activation kinetics. Moreover, as mentioned earlier, blocking the slow current in adult frog muscle does not abolish E-C coupling⁶. Further, both brief electrical stimuli and K-depolarization elicit vigorous mechanical responses of non-mutant myotubes and neonatal muscle fibres bathed in concentrations of cadmium sufficient to block both I_{fast} and I_{slow} (S. Jay and K.G.B., unpublished). Even though the actual current through the slow calcium channel seems unlikely to be directly involved in E-C coupling, the channel protein itself might still be intimately involved. For example, it has been suggested that the slow calcium-channel protein serves as the 'voltage-sensor' for the process governing SR calcium release²³. Accordingly, a mutation affecting slow

calcium current would also affect E-C coupling. It is also possible that the absence of I_{slow} and the failure of E-C coupling are only coincidentally related. The slow calcium channels in skeletal muscle are believed to be localized within the t-tubules^{4,24-26}, and in dysgenic muscle cells the normal morphological association between t-tubules and SR develops incompletely^{21,27}. Thus, this defect in t-tubular development may be the cause of both the absence of I_{slow} and the failure of E-C coupling.

We thank Dr Bruce Bean for many valuable discussions. This work was supported by NIH grant NS-14901 (K.G.B.) and grants from the Muscular Dystrophy Association of America (K.G.B. and J.A.P.), and the Blakeslee Fund of Smith College (J.A.P.). K.G.B. is the recipient of Research Career Development Award NS-00840.

Received 15 October 1985; accepted 20 January 1986.

- Costantin, L. L. *Prog. Biophys. molec. Biol.* **29**, 197-224 (1975).
- Endo, M. *Physiol. Rev.* **57**, 71-108 (1977).
- Sanchez, J. A. & Stefani, E. *J. Physiol., Lond.* **283**, 197-209 (1978).
- Almers, W., Fink, R. & Palade, P. T. *J. Physiol., Lond.* **312**, 177-207 (1981).
- Donaldson, P. L. & Beam, K. G. *J. gen. Physiol.* **82**, 449-468 (1983).
- Gonzalez-Serratos, H., Valle-Aguilera, R., Lathrop, D. A. & del Carmen Garcia, M. *Nature* **298**, 292-294 (1982).
- Gluecksohn-Waelsch, S. *Science* **142**, 1269-1276 (1963).
- Powell, J. A. & Fambrough, D. M. *J. cell. Physiol.* **82**, 21-38 (1973).
- Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85-100 (1981).
- Bean, B. P. *J. gen. Physiol.* **86**, 1-30 (1985).
- Nilius, B., Hess, P., Lansman, J. B. & Tsien, R. W. *Nature* **316**, 443-446 (1985).
- Bean, B. P., Sturek, M., Puga, A. & Hermesmeyer, K. *J. gen. Physiol.* **86**, 23a (1985).
- Armstrong, C. M. & Matteson, D. R. *Science* **277**, 65-67 (1985).
- Tsunoo, A., Yoshii, M. & Narahashi, T. *Biophys. J.* **47**, 433a (1985).
- Carbone, E. & Lux, H. D. *Nature* **310**, 501-502 (1984).
- Fedulova, S. A., Kostyuk, P. G. & Veselovsky, N. S. *J. Physiol., Lond.* **359**, 431-446 (1985).
- Bossu, J. L., Feltz, A. & Thomann, J. M. *Pflügers Arch. ges. Physiol.* **403**, 360-368 (1985).
- Nowicky, M. C., Fox, A. P. & Tsien, R. W. *Nature* **316**, 440-443 (1985).
- Bowden-Essien, F. *Dev Biol.* **27**, 351-364 (1972).
- Klaus, M. M., Stylianou, P. S., Rapalus, J. M., Briggs, R. T. & Powell, J. A. *Dev Biol.* **99**, 152-165 (1983).
- Pincon-Raymond, M., Rieger, F., Fosset, M. & Lazdunski, M. *Dev Biol.* **112**, 458-466 (1985).
- Haga, N. *et al. Cell* **39**, 71-78 (1984).
- Rios, E., Brum, G. & Stefani, E. *Biophys. J.* **49**, 13a (1986).
- Nicola Siri, L., Sanchez, J. A. & Stefani, E. *J. Physiol., Lond.* **305**, 87-96 (1980).
- Fosset, M., Jaimovich, E., Delpont, E. & Lazdunski, M. *J. biol. Chem.* **258**, 6086-6092 (1983).
- Glossman, H., Ferry, D. R. & Boschek, C. B. *Naunyn-Schmiedeberg Arch. Pharmacol.* **323**, 1-11 (1983).
- Powell, J. A. & Briggs, R. T. *J. Cell Biol.* **99**, 24a (1984).
- Bekoff, A. & Betz, W. J. *J. Physiol., Lond.* **271**, 25-40 (1977).
- Harary, I., Hoover, F. & Farley, B. *Meth. Enzym.* **32**, 740-745 (1974).

Spermatogenic failure in male mice lacking H-Y antigen

Paul S. Burgoyne, Elaine R. Levy & Anne McLaren

MRC Mammalian Development Unit, Wolfson House, 4 Stephenson Way, London NW1 2HE, UK

The mammalian Y chromosome carries a factor that initiates male sexual development by directing the fetal gonads to form testes. Wachtel and his colleagues¹ proposed that this testis-determining function of the Y is mediated by the male-specific cell-surface antigen H-Y, originally defined by skin grafting². This attractive hypothesis, which has been widely accepted, was based on the assumption that serological tests using antisera raised against male cells were recognizing H-Y antigen. Although disputed³, this assumption is supported by some recent studies⁴⁻⁶. However, mice have been described⁷ which develop testes but lack the cell-surface H-Y antigen as defined by T-cell-mediated transplantation tests. Thus, although it remains possible that a serologically detected male-specific antigen is responsible for testis determination, it seems that H-Y, as originally defined, is not. We show here that H-Y-negative male mice, in losing the genetic information that encodes H-Y, have also lost genetic information required for spermatogenesis. This result identifies a gene on the mouse Y,

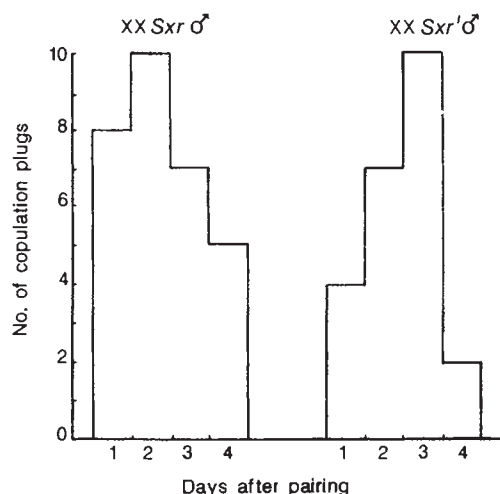


Fig. 1 Copulatory activity of XX *Sxr'* compared with XX *Sxr* males. Four males of each type were each paired with three females (MF1 strain; Olac) and the females were examined for copulation plugs on the following four mornings. This was repeated for four successive weeks. The histograms show the total numbers of plugs found on days 1–4 after pairing.

distinct from the testis-determining gene, which is necessary for spermatogenesis, and raises the intriguing possibility that the product of this 'spermatogenesis gene' is H-Y antigen.

Cattanach *et al.*⁸ originally described a sex reversal factor (*Sxr*) in mice which caused XX individuals to develop as males with testes. It is now known that *Sxr* is a small segment of the Y chromosome which is transferred to the X during meiosis in XY *Sxr* carriers^{9,10}. *Sxr* includes the genetic information for testis determination and for the expression of H-Y¹¹. However, McLaren *et al.*⁷ found a variant *Sxr* (designated *Sxr'*) that still carried the testis-determining information, but did not confer H-Y antigenicity. Our objective here was to determine whether loss of H-Y antigenicity was correlated with any impairment of male reproductive function, as it seemed likely that this male-specific antigen would have a male-specific role. Initially, we compared XX *Sxr'* males with XX *Sxr* males. The H-Y-negative XX *Sxr'* males were not detectably less proficient than XX *Sxr* males at mating or at inducing synchronization of mating behaviour (Fig. 1), suggesting that testosterone and pheromone production are unaltered in the absence of H-Y antigen.

Adult XX *Sxr'* mice, like XX *Sxr* mice, have sterile testes due to the perinatal loss of germ cells (E.R.L. and P.S.B., unpublished observations), characteristic of male mice carrying two X chromosomes. Presumably the expression of a double dose of X-linked genes in the germ cells is incompatible with their survival beyond the perinatal period. Adult X0 *Sxr* mice, in contrast to XX *Sxr'* and XX *Sxr* mice, have all stages of spermatogenesis represented in their testes. They are nonetheless sterile because the later stages are defective, and only a few grossly abnormal sperm are produced⁸. These defects may result from the absence of a pairing partner for the X chromosome, since X-Y pairing during pachytene seems to be necessary for the successful completion of spermatogenesis¹².

We produced X0 *Sxr'* mice and compared them with X0 *Sxr* mice to see if the loss of H-Y antigenicity is correlated with any additional defect of spermatogenesis. One X0 *Sxr* and four X0 *Sxr'* mice were checked for their H-Y status (Table 1). For each mouse designated H-Y-negative (final column), the target cells were shown to be negative with appropriate H-2-restricted H-Y-specific cytotoxic T cells but positive with the corresponding anti-H-2 cytotoxic T cells. Each of the X0 *Sxr'* mice was H-Y-negative by these criteria. By contrast, the X0 *Sxr* male typed positive, like the XY control male and the previously

Table 1 H-Y typing of X0 *Sxr'*, X0 *Sxr* and control XX and XY mice with H-2^k and H-2^d-restricted H-Y-specific T cells

Mouse	H-2†	Karyotype	Sex	% Specific lysis at E:T = 10:1 T _c *		H-Y summary
				Anti-H-2 ^q	Anti-H-Y ^q	
555	bq	X0 <i>Sxr'</i>	Male	28	1	—
556	bq	X0 <i>Sxr'</i>	Male	27	0	—
SWR	qq	XY	Male	38	46	+
SWR	qq	XX	Female	31	9	—
				Anti-H-2 ^k	Anti-H-Y ^k	
557	kk	X0 <i>Sxr</i>	Male	20	22	+
558	kb	X0 <i>Sxr'</i>	Male	17	1	—
559	kb	X0 <i>Sxr'</i>	Male	29	6	—
CBA	kk	XY	Male	29	38	+
CBA	kk	XX	Female	38	9	—

* Cytotoxic T cells (T_c) were from 5-day mixed lymphocyte cultures (MLC) of (B10×BALB/c)F₁ anti-SWR (anti-H-2^q), (B10×SWR)F₁ anti-SWR (anti-H-Y^q), BALB.B anti-CBA (anti-H-2^k) and (B10×CBA)F₁ anti-CBA (anti-H-Y^k). The anti-H-2 were primary, the anti-H-Y secondary *in vitro* MLC and were used as attacker T_c as described previously⁷ with 5-day concanavalin A-stimulated spleen cells from the mice to be typed as ⁵¹Cr-labelled targets. Four effector/target (E:T) cell ratios were tested for each target using serial threefold dilutions of the effector T_c starting at E:T = 30:1, and a fixed number (10⁶) of target cells. The data were analysed by regression analysis; the values given are from the titration curves, at E:T = 10:1. Numbers underlined are positive values, and lie on curves with *r*² > 0.80 (*r* = coefficient of correlation, a measure of goodness of fit of the data on the regression line).

† H-2 typing with monoclonal antibody plus rabbit complement, as described previously⁷.

reported XX *Sxr* males¹³, with both H-2-restricted H-Y-specific cytotoxic T cells and the corresponding anti-H-2 cytotoxic T cells. Histological analysis revealed that spermatogenesis is much more severely affected in X0 *Sxr'* than X0 *Sxr* mice. There is almost total elimination of spermatogenic cells beyond the spermatogonial stage in adult X0 *Sxr'* testes, and this block is evident prepubertally, at the onset of meiosis (Fig. 2). Although grossly the block to spermatogenesis is clearcut, it is hard to define the point at which it acts. Spermatogonia appear to be few in number and spermatogonial divisions are rare: this could result from either an intrinsic defect of proliferation in the spermatogonial compartment, or some inhibition of proliferation in response to a block in spermatogonial differentiation and commitment to meiosis. Even the block to meiosis is not absolute, because very occasionally patches of cells enter meiosis (Fig. 3), only to arrest and degenerate early in the pachytene stage.

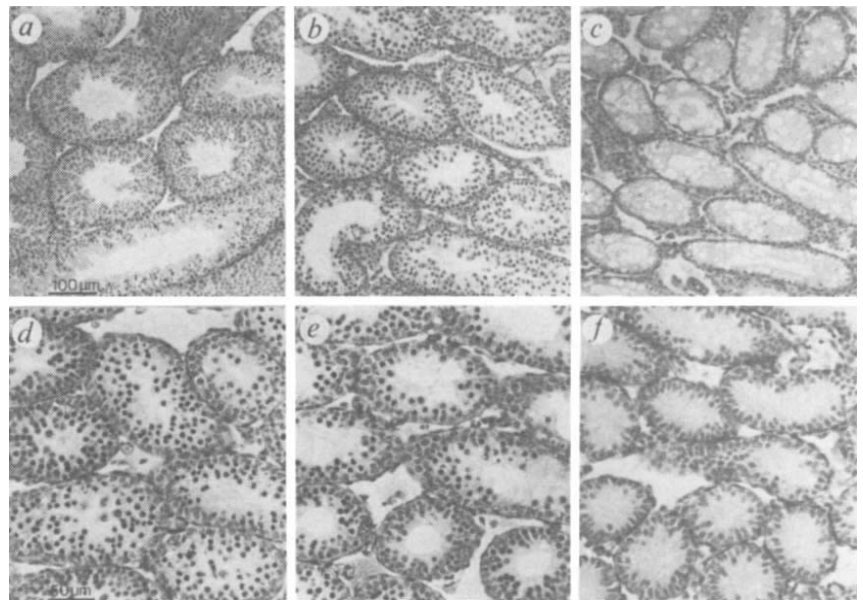
Our results, together with those of McLaren *et al.*⁷, show that when *Sxr'* arose from *Sxr*, genetic information needed for H-Y expression and spermatogenesis was lost, while genetic information required for testis determination was retained. Incidentally, the fact that X0 *Sxr'* mice are H-Y-negative, just like T16H/X *Sxr'* mice⁷, rules out Ohno's recent suggestion that spreading of X-inactivation into *Sxr'* is responsible for the non-expression of H-Y¹⁴.

Levy and Burgoyne¹⁵ found that there is an early spermatogenic failure of X0 germ cells in X0/XY mosaic mice, and concluded that since the X0 germ cells are not 'rescued' by the mosaic testicular soma, there must be a 'spermatogenesis gene' on the Y chromosome which is expressed in the germ line. Histologically, the arrested spermatogenesis that occurs in patches in these X0/XY mosaic testes is indistinguishable from that in X0 *Sxr'* testes; this suggests that it is the loss or mutation of this 'spermatogenesis gene' which leads to the spermatogenic arrest in X0 *Sxr'* mice.

As the loss of spermatogenesis in X0 *Sxr'* mice is correlated with the loss of H-Y expression, it is tempting to conclude that H-Y antigen is the product of the spermatogenesis gene, thus reinstating a male-specific function for this evolutionarily conserved male antigen. The same correlation of H-Y antigen negativity and spermatogenic failure was seen in the X0 male mouse described by Melvold *et al.*¹⁶. However, it remains possible that there are separate genes for spermatogenesis and H-Y

Fig. 2 Spermatogenic failure in X0 *Sxr'* male mice. *a*, Complete spermatogenesis in a testis from an adult (49-day-old) XY mouse. *b*, Abnormal spermatogenesis with a marked deficiency of condensing spermatid stages in the testis of a 49-day-old X0 *Sxr* mouse (H-Y-positive). *c*, An almost total block of spermatogenesis in the testis of a 49-day-old X0 *Sxr'* mouse (H-Y-negative). Although some spermatogonia are present, almost all tubules lack spermatocytes and spermatids. *d-f*, Testes from 13-day-old XY, X0 *Sxr* and X0 *Sxr'* mice showing that the defect in X0 *Sxr'* spermatogenesis is already present prepubertally when meiosis begins.

Methods. X0 *Sxr* and X0 *Sxr'* mice were produced by mating XY *Sxr* or XY *Sxr'* males to females heterozygous for the inversion In(X)1H. These females generate some nullo-X eggs following crossing-over within the inversion. The testes from 15 X0 *Sxr'* and 9 X0 *Sxr* males together with those from their male littermates, at 2–50 days post partum, were examined during this study. All mice were karyotyped. Testes were fixed in Bouin's fixative. The sections were stained with periodic acid-Schiff's reagent and counterstained with haematoxylin.



expression which are tightly linked. It may be pertinent that spermatogonia have been identified as the only adult male cells which type negative for male-specific antigen using serological tests¹⁷. Later spermatogenic stages are increasingly positive. If serological tests do recognize H-Y antigen, this pattern of expression would be consistent with H-Y antigen being the spermatogenesis gene product.

The possibility that H-Y antigen is the product of the spermatogenesis gene leads to predictions about the H-Y status of human XX males and XY females. A human spermatogenesis gene, which we presume is homologous to that on the mouse Y, is located on the human Y long arm close to the centromere¹⁸. Human XX males that have acquired testis-determining sequences from the Y short arm^{19–21} will lack Y long arm sequences including the spermatogenesis gene, and hence should type H-Y-negative if H-Y is the spermatogenesis gene product. Conversely, XY females that have lost testis-determining sequences from the Y short arm should type H-Y-positive because the spermatogenesis gene on the Y long arm should have been retained. H-Y typing of human sex-reversed individuals is now possible using histocompatibility locus antigen (HLA)-restricted H-Y-specific T-cell clones. All five XX males so far typed are H-Y-negative, in line with our predictions (ref. 22 and E. Goulmy and E. Simpson, personal communication).

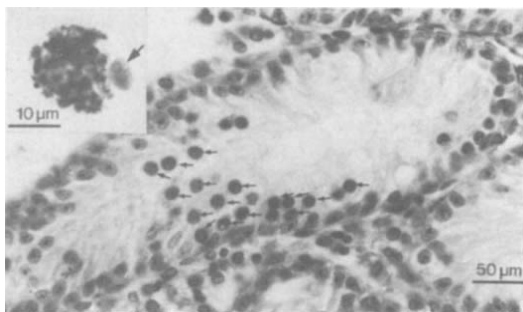


Fig. 3 Meiotic cells in X0 *Sxr'* testes. Occasionally tubules are seen in X0 *Sxr'* testes which contain patches of meiotic cells (arrowed) that degenerate during the early pachytene stage. These pachytene cells have a clear sex vesicle in air-dried preparations (arrowed in inset), showing that this structure can form in the absence of H-Y antigen.

We thank Dr Elizabeth Simpson of the Clinical Research Centre, Harrow, for H-Y typing. E.R.L. is a recipient of an SERC studentship.

Received 20 November 1985; accepted 29 January 1986.

- Wachtel, S., Ohno, S., Koo, G. C. & Boyse, E. A. *Nature* **257**, 235–236 (1975).
- Eichwald, E. J. & Silmsker, C. R. *Transplant Bull.* **2**, 154–155 (1955).
- Silvers, W. K., Gasser, D. L. & Eicher, E. M. *Cell* **28**, 439–440 (1982).
- Wiberg, U. H. *Hum. Genet.* **69**, 15–18 (1985).
- Wiberg, U. H. & Günther, E. *Immunogenetics* **21**, 91–96 (1985).
- Wiberg, U. H. & Mayerova, A. *J. Immunogenet.* **12**, 55–63 (1985).
- McLaren, A., Simpson, E., Tomonari, K., Chandler, P. & Hogg, H. *Nature* **312**, 552–555 (1984).
- Cattanach, B. M., Pollard, C. E. & Hawkes, S. G. *Cytogenetics* **10**, 318–337 (1971).
- Singh, L. & Jones, K. W. *Cell* **28**, 205–216 (1982).
- Evans, E. P., Burtenshaw, M. D. & Cattanach, B. M. *Nature* **300**, 443–445 (1982).
- Bennett, D. *et al. Nature* **265**, 255–257 (1977).
- Burgoyne, P. S. & Baker, T. G. in *Controlling Events in Meiosis* (eds Evans, C. W. & Dickinson, H. G.) 349–362 (Company of Biologists, Cambridge, 1984).
- Simpson, E., Edwards, P., Wachtel, S., McLaren, A. & Chandler, P. *Immunogenetics* **13**, 355–358 (1981).
- Ohno, S. *Endocrine Rev.* **6**, 421–431 (1985).
- Levy, E. R. & Burgoyne, P. S. *Cytogenet. Cell Genet.* (submitted).
- Melvoid, R. W., Koln, H. I., Yerganian, G. & Fawcett, D. W. *Immunogenetics* **5**, 33–41 (1977).
- Zenzen, M. T., Müller, U., Aschmoneit, I. & Wolf, U. *Hum. Genet.* **45**, 297–303 (1978).
- Tiepolo, L. & Zuffardi, O. *Hum. Genet.* **34**, 119–124 (1976).
- de la Chapelle, A., Tippet, P. A., Wetterstrand, G. & Page, D. *Nature* **307**, 170–171 (1984).
- Guellaen, G. *et al. Nature* **307**, 172–173 (1984).
- Page, D. C., de la Chapelle, A. & Weissenbach, J. *Nature* **315**, 224–226 (1985).
- Goulmy, E., van Leeuwen, A., Blokland, E., Sachs, E. S. & Geraedts, J. P. M. *Immunogenetics* **17**, 523–531 (1983).

Modulation of visual cortical plasticity by acetylcholine and noradrenaline

Mark F. Bear* & Wolf Singer

Max Planck Institute for Brain Research, Deutschordenstrasse 46, 6000 Frankfurt 71, FRG

During a critical period of postnatal development, the temporary closure of one eye in kittens will permanently shift the ocular dominance (OD) of neurones in the striate cortex to the eye that remains open¹. The OD plasticity can be substantially reduced if the cortex is infused continuously with the catecholamine neurotoxin 6-hydroxydopamine (6-OHDA) during the period of monocular deprivation^{2–5}, an effect that has been attributed to selective depletion of cortical noradrenaline⁶. However, several

* Present address: Center for Neural Science, Box 1953, Brown University, Providence, Rhode Island 02912, USA.

Fig. 1 *a, b*, Coronal sections through the forebrain of a 44-day-old kitten (K125) at the approximate frontal planes A14.0 and A11.75, respectively. Projected onto the right side of *a* and *b* is the distribution of cells in the basal telencephalon that were back-filled after 1 μ l of 30% horseradish peroxidase (HRP) was injected into the visual cortex of a 4-week-old kitten. For each drawing, five contiguous 50- μ m thick sections were pooled; each dot represents approximately two HRP-labelled neurones. (????) The extrinsic cholinergic innervation of the striate cortex arises from these cells²⁰. This unilateral projection derives from neurones scattered within the internal capsule (IC), the substantia innominata (SI), the diagonal band of Broca and the medial septal nucleus (SMN). Reconstructed on the left side of *a* and *b* are the regions of cell loss (crosshatching) after NMA was injected into the telencephalon of K125. Also indicated (●—) are the two needle tracks along which NMA was delivered, one (*a*) targeting cells in the vertical limb of the diagonal band (50 μ g) and SMN (50 μ g), the other (*b*) targeting cells in the horizontal limb of the diagonal band (DBH, 125 μ g), SI (75 μ g) and IC (50 μ g). This lesion produced a drastic and widespread depletion of AChE-positive axons in the neocortex. In addition to the cholinergic basal forebrain neurones, cells in the caudate, globus pallidus, ventral pallidum and rostral pole of the dorsal thalamus were also destroyed by the NMA injections. The cell loss in the thalamus was confined to the reticular, ventral anterior and ventral lateral nuclei. The portions of the claustrum (Clau) and intralaminar thalamus that project to the visual cortex were spared by the NMA injections. AC, anterior commissure; Ca, caudate nucleus. *c*, Coronal section through the forebrain of K152 at the approximate frontal plane A16.25 to illustrate a cingulate gyrus lesion. The blackened region was surgically removed by subpial aspiration. The lesion produced a depletion of AChE-positive axons in the striate cortex comparable to that observed with basal forebrain lesions. Analysis of cortical tissue using HPLC confirmed that NA was also reduced in area 17 to <50% of control levels (Table 1). *d*, Anterior-posterior extent of the lesion in *c* in a mid-sagittal view of a kitten brain, showing the approximate level of each of the sections illustrated. Also indicated is the extent of the cingulate gyrus lesion in K152 (black area). Scale bars, 5.0 mm.

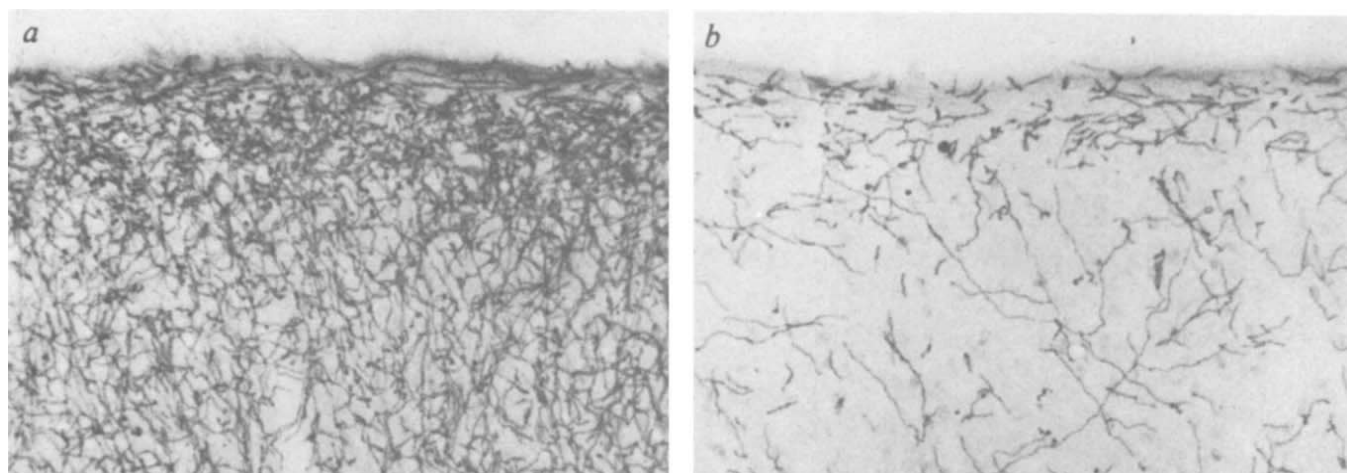
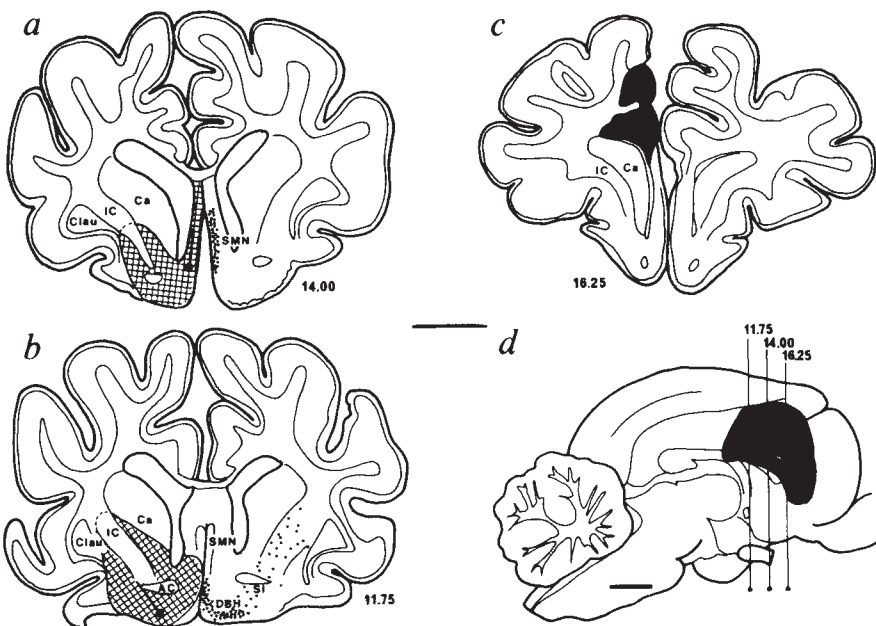


Fig. 2 Photomicrographs of AChE-containing axons in the striate cortex of a 47-day-old kitten. *b*, Layers I-IV in the hemisphere that had received a chemical lesion of the cholinergic basal telencephalon 18 days earlier. *a*, A homotypic region in the contralateral control hemisphere. A unilateral loss of ocular dominance plasticity accompanies this reduction in the density of cholinergic axons if cortical noradrenaline is depleted concurrently. Magnification, $\times 150$.

other methods causing noradrenaline (NA) depletion leave the plasticity intact⁷⁻¹⁰. Here we present a possible explanation for the conflicting results. Combined destruction of the cortical noradrenergic and cholinergic innervations reduces the physiological response to monocular deprivation although lesions of either system alone are ineffective. We also find that 6-OHDA can interfere directly with the action of acetylcholine (ACh) on cortical neurones. Taken together, our results suggest that intracortical 6-OHDA disrupts plasticity by interfering with both cholinergic and noradrenergic transmission and raise the possibility that ACh and NA facilitate synaptic modifications in the striate cortex by a common molecular mechanism.

Our study was designed to assess the contribution of the extrathalamic cortical afferents to OD plasticity in the striate

cortex. Our attention was focused initially on the cholinergic projection for three reasons. First, there is converging biochemical^{11,12} and anatomical¹³ evidence suggesting that the striate cortex receives a dense cholinergic innervation during the critical period. Second, the facilitatory action of ACh on excitatory transmission in striate cortex^{14,15} increases the probability of postsynaptic activation, a condition that is required for the experience-dependent modification of many excitatory synapses¹⁶. Third, the cholinergic projection is thought to be an important component of the ascending reticular activating system¹⁷, and there is evidence that reticular activation of striate cortex may be necessary for OD plasticity^{18,19}.

The cholinergic innervation of striate cortex arises from cells scattered through the basal telencephalon (Fig. 1*a, b*; ref. 20).

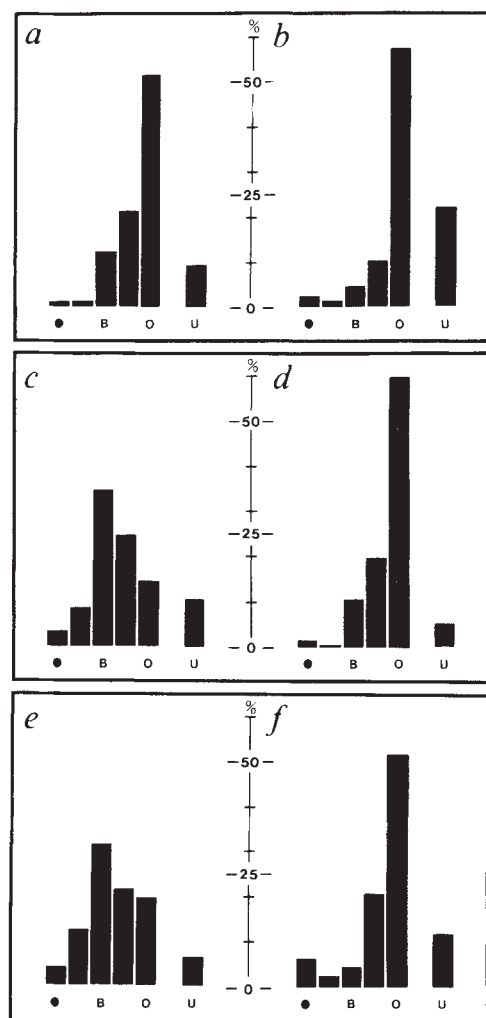


Fig. 3 Percentages of cells in each of the five ocular dominance groups in the experimental (*a, c, e*) and control (*b, d, f*) hemispheres. Open circles, monocular open eye groups; filled circles, monocular closed eye groups; B, strictly binocular groups (group 3); U, % of unclassifiable neurones. *a, b*, Composite ocular dominance histograms from kittens that received unilateral NMA lesions of the cholinergic basal forebrain before 7–11 days of monocular deprivation (see Table 1). The ocular dominance distribution is shifted to the open eye both in the ACh-depleted left hemispheres (*a*, 126 cells) and in the control right hemispheres (*b*, 66 cells). *c, d*, Composite OD histograms from kittens that received unilateral 6-OHDA lesions of the dorsal noradrenergic bundle 1 week before the NMA lesions of the forebrain (see Table 1). *c*, OD distribution in the experimental hemispheres (110 cells) is significantly more binocular and less shifted than in *d*, showing the distribution in the control hemispheres (97 cells). *e, f*, Composite OD histograms from kittens that received unilateral lesions of the cingulate gyrus before 8–9 days of monocular deprivation (see Table 1). The OD histogram from the cingulate-damaged hemispheres (*e*, 94 cells) is also significantly more binocular and less shifted than that from the control hemispheres (*f*, 64 cells). The striate cortex in both the 6-OHDA + NMA-injected and cingulate-damaged hemispheres was depleted of its normal complement of cholinergic and noradrenergic axons. Other than OD, the visual-response properties of the cells recorded in these hemispheres did not differ from the controls. Histological reconstruction of the electrode tracks was not possible because the area centralis representation of area 17 was dissected for HPLC analysis. However, our sampling methods were the same in both hemispheres and no obvious laminar bias was indicated by the relative percentages of simple or complex cells encountered.

Table 1 Rearing histories of kittens and physiological results

Animal	Lesion	Age at lesion (days postnatal)	Monocular deprivation (days postnatal)	Control hemispheres			Experimental hemispheres			% NA
				N	B	OED	N	B	OED	
K125	BF	29	33–43	21	0.06	0.94	30	0.46	0.61	—
K127	BF	29	40–47	20	0.25	0.81	30	0.38	0.77	—
K128	BF	29	44–55	25	0.33	0.67	36	0.52	0.59	—
K142	BF	32	32–40	—	—	—	30	0.26	0.83	—
K149	DNAB/BF	32/39	39–48	30	0.39	0.71	35	0.84	0.31	16
K151	DNAB/BF	32/39	41–50	30	0.33	0.80	35	0.78	0.39	49
K156	DNAB/BF	32/39	39–49	37	0.31	0.72	40	0.74	0.26	40
K152	C	39	39–47	31	0.19	0.89	31	0.69	0.40	30
K157	C	39	39–47	—	—	—	33	0.70	0.32	40
K158	C	39	39–48	33	0.42	0.60	30	0.79	0.30	31
K160	DNAB	32	39–49	—	—	—	30	0.33	0.82	33
K164	C*	35	35–45	—	—	—	30	0.33	0.64	—

BF, basal forebrain; DNAB, dorsal noradrenergic bundle; DNAB/BF, DNAB followed by BF; C, surgical aspiration of the cingulate gyrus; C*, NMA lesion of cingulate cortex. N, number of cells recorded in each hemisphere; B, binocularity; OED, open-eye dominance for each hemisphere. Binocularity is defined as the number of cells in OD groups 2–4 divided by the total number of cells recorded. OED (after Paradiso *et al.*²⁸) is defined as the number of cells in group 5 plus 0.5 times the number of cells in group 4 divided by the total number of cells (group 5, by convention, is the monocular group dominated by the open eye). In all cases except K142 the lesions were unilateral so that the opposite hemispheres could serve as controls. The cortical NA content in hemispheres with DNAB and C lesions (expressed as a per cent of the control cortex) is also shown. In these cases, saline-perfused samples of visual cortex were rapidly dissected from both hemispheres and frozen at -70°C . The samples were subsequently homogenized in 0.1 M perchloric acid containing EDTA (60 mg l^{-1}) and sodium metabisulphite (100 mg l^{-1}) and then centrifuged to remove denatured protein. Aliquots of the supernatant were applied directly to an HPLC column (ALTEX ODS $5\text{ }\mu\text{m}$) and the catechol compounds were measured relative to an internal dihydroxybenzylamine standard using electrochemical detection as described elsewhere^{7,28}. The NA content of the control visual cortex was 73.67 ± 7.26 (mean $\pm$ s.e.m.) ng per g tissue, wet weight.

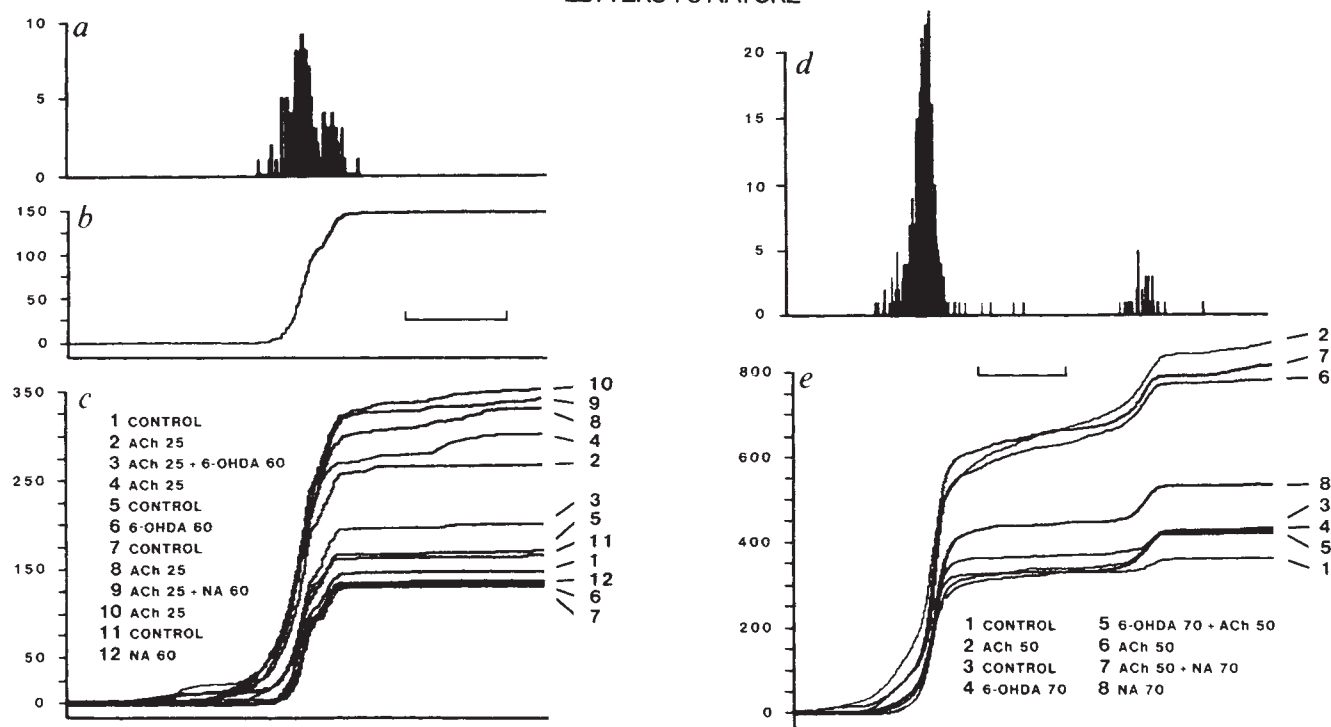


Fig. 4 Antagonistic effect of iontophoretically applied 6-OHDA on ACh-evoked responses. *a*, Peri-stimulus time histogram of the control response of a simple cell in the striate cortex to an optimally oriented bar of light swept at a constant speed over the receptive field of the neurone. The ordinate shows the number of spikes per bin accumulated in 10 trials; bin width, 20 ms. *b*, To facilitate quantitative comparisons, the same data are expressed as total spikes accumulated as a function of sweep time. In this example, the cell responded to 10 stimulus presentations with 148 spikes. Scale bar, 1.0 s. *c*, Response of this cell to visual stimulation is plotted as in *b*, under the 12 test conditions listed in chronological order. The interval between tests ranged from 1 to 1.5 min. Indicated after each drug tested is the iontophoresis current (nA). The visual response of this cell was reliably enhanced each time ACh was applied by itself (tests 2, 4, 8 and 10). However, this enhancement was reduced considerably when 6-OHDA was applied concurrently (test 3). In contrast, NA applied in the same way had no effect on the ACh response (test 9); 6-OHDA application by itself had little apparent effect (test 6) compared with the control responses (tests 1, 5, 7 and 11). *d*, Peri-stimulus time histogram of the control response of a complex cell in the striate cortex to a bar of light swept first in the preferred direction, then in the opposite direction. *e*, Data from this complex cell expressed as in *b* and *c*. Test conditions are listed in chronological order and the interval between tests ranged from 1 to 7 min. In this example, the facilitation of the visual response by ACh (tests 2 and 6) was completely blocked by 6-OHDA (test 5). Noradrenaline did not have a similar action (test 7) and 6-OHDA had no effect on the cellular response when applied alone (test 4). Scale bar, 1.0 s. Excitation of this cell with glutamate was not antagonized by the concurrent application of 6-OHDA (data not shown).

Methods. All the records were made with a micropipette containing 1.5 M potassium citrate ($R = 15 \text{ M}$) that protruded $\sim 50 \mu\text{m}$ from an attached 7-barrel micropipette. Two of these barrels contained acetylcholine chloride (2.0 M, pH 4.5), two contained 6-OHDA-HCl (0.2 M) in 0.1% ascorbate (pH 3), one was filled with noradrenaline-HCl (0.2 M), also in 0.1% ascorbate, and one contained glutamic acid (0.3 M, pH 8). Currents of 5 nA were applied to retain the drugs in the barrels; positive for glutamate and negative for all others. Tip potentials were automatically balanced by passing currents of equal and opposite polarity through an additional barrel containing 2.0 M NaCl.

We destroyed these cells with the excitotoxin *N*-methyl-DL-aspartate (NMA) in 4-5-week-old kittens, and then tested for a deficit in the OD shift after brief periods of monocular deprivation. All lesions were made in the left hemisphere. The right eyelids were closed and, 7-10 days later, a standard physiological assay of OD was performed in the striate cortex on both sides¹⁸. As the normal OD shift is nearly symmetrical in the two hemispheres^{8,9}, the cortex on the right (unoperated) side served as an internal control. The effectiveness of the lesions was monitored by acetylcholinesterase (AChE) histochemistry. Most, if not all, of the AChE-positive axons in the cat striate cortex arise from cholinergic cells in the basal telencephalon²⁰.

A typical basal forebrain (BF) lesion, illustrated in Fig. 1*a, b* (left side), resulted in a drastic reduction in the density of AChE-stained axons in striate cortex (Fig. 2). Nevertheless, subsequent monocular deprivation shifted the OD of cortical neurones dramatically to the open eye (Fig. 3*a, b*, Table 1). These data indicate that a substantial depletion of ACh alone is not sufficient to block plasticity. However, when cortical NA was depleted by $>50\%$ (Table 1) before the basal forebrain lesions, the normal OD shift was prevented (Fig. 3*c, d*; Table 1). The noradrenergic projection was destroyed unilaterally in these animals by injecting 6-10 μg of 6-OHDA into the left dorsal noradrenergic bundle (co-ordinates A7, L3, DO). Daw

*et al.*⁹ have shown, and we have confirmed (K160, Table 1), that dorsal noradrenergic bundle lesions do not normally affect the plastic response to monocular deprivation, even with a 70-90% depletion of cortical NA. Thus, although basal forebrain and dorsal noradrenergic bundle lesions alone are ineffective, they cause a significant loss of OD plasticity in the striate cortex when they are combined.

These data suggest that the combined depletion of cortical ACh and NA is a sufficient condition to retard experience-dependent synaptic modifications in the striate cortex. This notion was tested further by using a second experimental approach. Anatomical studies in this laboratory indicate that cholinergic and noradrenergic axons en route to area 17 travel together within the white matter of the cingulate gyrus. Thus, surgical lesions of the cingulate gyrus also effectively deplete striate cortex of both NA and ACh. Such a lesion (Fig. 1*c, d*) drastically reduces the density of AChE-positive axons, depletes endogenous NA (Table 1) and prevents the normal OD shift after monocular deprivation in area 17 on the lesioned side (Fig. 3*e, f*; Table 1). In one kitten (Table 1, K164), an extensive lesion of the cingulate cortex was made with multiple injections of NMA, which spares axons of passage. This treatment did not significantly alter the normal OD shift after monocular deprivation, thus confirming that the effects of the surgical lesions are

caused by the interruption of pathways passing through the cingulate gyrus.

These results predict that any treatment that simultaneously blocks noradrenergic and cholinergic transmission in the cortex is effective in slowing the OD shift after monocular deprivation. The work of Furness²¹ on intestinal smooth muscle indicates that continuously applied 6-OHDA has exactly this effect—noradrenergic denervation and blockade of muscarinic cholinergic transmission. We explored this possibility in the striate cortex by assessing the effects of 6-OHDA iontophoresis on the neuronal responses induced by applying ACh from a piggyback microelectrode assembly²². We studied 30 ACh-sensitive neurones in the striate cortex of three adult cats. The effects of ACh iontophoresis were generally in agreement with those of Sillito and Kemp¹⁴. In most cases (26 cells), the visual responses were enhanced by ACh. However, in 85% of these cells the enhancement was significantly attenuated when ACh was applied concurrently with 6-OHDA (Fig. 4).

Our iontophoretic data are consistent with the hypothesis that the continuous application of 6-OHDA to the cortex affects cholinergic transmission, suggesting that intracortical 6-OHDA disrupts cortical plasticity by a combined action on the noradrenergic and cholinergic projections. In this context, note that the threshold concentration of 6-OHDA to block muscarinic transmission in the gut is $\sim 5 \mu\text{M}$ (ref. 21). This is the same tissue concentration of 6-OHDA that Kasamatsu *et al.*²³ calculated to be the threshold for preventing plasticity in the striate cortex of kittens.

It is not yet clear whether these data indicate a special role for NA and ACh in the control of cortical plasticity. Many experience-dependent synaptic modifications seem to require the activation of cortical neurones¹⁶ and there are indications that this activation must exceed the threshold of voltage-sensitive Ca^{2+} channels in order to be effective²⁴. Thus, removal of enough facilitatory extrageniculate inputs could lower cortical excitability below the threshold for synaptic modifications. Both ACh and NA reduce K^+ permeability^{25,26} and hence both could enhance depolarization in response to visual input¹⁵. On the other hand, the actions of these neuromodulators may be related more specifically to the control of synaptic plasticity in the cerebral cortex. For example, both stimulate the production of cyclic nucleotides and the mobilization of intracellular Ca^{2+} in target neurones²⁷. Thus, it is plausible that ACh and NA regulate, via second-messenger-dependent phosphorylation, a common set of proteins that are involved in the modification of synaptic transmission.

We thank Dr P. Zöphel for performing the HPLC analysis, Dr J. Dann for critically reading the manuscript, C. Ziegler for making the electrodes, E. Kastl and A. Franke for assistance with the histology, and G. Trauten for typing the manuscript.

Received 27 June; accepted 12 November 1985.

- Wiesel, T. N. & Hubel, D. H. *J. Neurophysiol.* **26**, 1003–1017 (1963).
- Kasamatsu, T. & Pettigrew, J. D. *Science* **194**, 206–209 (1976).
- Kasamatsu, T. & Pettigrew, J. D. *J. comp. Neurol.* **185**, 139–162 (1979).
- Kasamatsu, T., Pettigrew, J. D. & Ary, M. J. *comp. Neurol.* **185**, 163–182 (1979).
- Pettigrew, J. D. & Kasamatsu, T. *Nature* **271**, 761–763 (1978).
- Kasamatsu, T. *Prog. Psychobiol. physiol. Psychol.* **10**, 1–112 (1984).
- Bear, M. F. *et al. Nature* **302**, 245–247 (1983).
- Daw, N. W., Robertson, T. W., Rader, R. K., Videen, T. O. & Coscia, C. J. *J. Neurosci.* **4**, 1354–1360 (1984).
- Daw, N. W., Videen, T. O., Parkinson, D. & Rader, R. K. *J. Neurosci.* **5**, 1925–1933 (1985).
- Adrien, J. *et al. J. Physiol., Lond.* **367**, 73–98 (1985).
- Potempska, A., Skangiel-Kramska, J. & Kossut, M. *Dev. Neurosci.* **2**, 38–45 (1979).
- Shaw, C., Needler, M. C. & Cynader, M. *Dev. Brain Res.* **14**, 295–299 (1984).
- Bear, M. F., Carnes, K. M. & Ebner, F. F. *J. comp. Neurol.* **237**, 519–532 (1985).
- Sillito, A. M. & Kemp, J. A. *Brain Res.* **289**, 143–155 (1983).
- Sillito, A. M. *Nature* **303**, 477–478 (1983).
- Rauschecker, J. P. & Singer, W. *Nature* **280**, 58–60 (1979).
- Singer, W. in *The Neurosciences: 4th Study Program*, 1093–1110 (MIT Press Cambridge, 1979).
- Singer, W. *Expl Brain Res.* **47**, 209–222 (1982).
- Singer, W. & Rauschecker, J. P. *Expl Brain Res.* **47**, 223–233 (1982).
- Bear, M. F., Carnes, K. M. & Ebner, F. F. *J. comp. Neurol.* **234**, 411–430 (1985).
- Furness, J. B. in *6-Hydroxydopamine and Catecholamine Neurons*, 205–214 (North-Holland, Amsterdam, 1971).
- Francesconi, W., Müller, C. M. & Singer, W. *Neurosci. Lett.* **18**, 309 (1984).
- Kasamatsu, T., Itakura, T. & Jonsson, G. *J. Pharmac. exp. Ther.* **217**, 841–850 (1981).
- Geiger, H. & Singer, W. *Expl Brain Res. Suppl.* (in the press).
- Madison, D. V. & Nicoll, R. A. *Nature* **299**, 636 (1982).
- Halliwel, J. V. & Adams, P. R. *Brain Res.* **250**, 71 (1982).
- Nestler, E. J., Walaas, S. I. & Greengard, P. *Science* **225**, 1357–1364 (1984).
- Paradiso, M. A., Bear, M. F. & Daniels, J. D. *Expl Brain Res.* **51**, 413–422 (1983).

T-cell recognition of antigen and the Ia molecule as a ternary complex

Jonathan D. Ashwell & Ronald H. Schwartz

Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA

T-lymphocyte co-recognition of antigen and major histocompatibility complex (MHC)-encoded molecules (such as murine Ia molecules) is thought to be mediated by a single cell-surface receptor, although the molecular mechanism by which this occurs is controversial (reviewed in ref. 1). One possibility is that the antigen molecule and the Ia molecule interact physically, either before or after encountering the T-cell antigen-specific receptor^{1,2}. Alternatively, both molecules could bind to the receptor independently of one another, accounting for the dual specificity of the receptor without postulating a physical interaction between a limited number of Ia molecules present in any given animal and the myriad antigens to which T cells can respond. Here, we used a recently described approach for analysing the relative avidity of the T-cell receptor for different ligands³ to address these two possibilities. We describe a T-cell clone whose response to a single antigen, presented in the context of two different Ia molecules, strongly suggests that the antigen and the Ia molecule interact physically.

Receptor-bearing cells in culture compete with each other for ligand. For a fixed concentration of ligand, as the number of responding cells in an assay increases, at some point the number of receptors occupied per cell will decrease as the free ligand becomes significantly depleted by binding to the receptor. If the response of the cell to the ligand is proportional to the number of receptors occupied, at least in the dose-sensitive portion of the response, then the average cellular response, or the probability that a cell will respond in the case of a proliferation assay, will also decline^{4–6}. Furthermore, the point at which the cells bind sufficient ligand to result in a decrease in the average cellular response is dependent upon two variables: the number of receptors per cell and the affinity of the receptors for the ligand. We have used this phenomenon to analyse the interaction of the T-cell antigen-specific receptor with its ligand, which is a combination of antigen and Ia molecule (antigen-Ia)³. The fraction of T cells stimulated by antigen-Ia to incorporate ³H-thymidine varied as a function of the number of responding T cells, best displayed by plotting the response of the T cells as a fraction of the maximum response achieved (Fig. 1). In a previous study we used this technique to compare the ability of four synthetic analogues of moth cytochrome *c* to activate a T-cell clone when presented in association with a single Ia molecule, $E_{\beta}^k E_{\alpha}^k$ (ref. 3). In the present study, we assessed the effect of changing the allelic form of the Ia molecule while keeping the antigen constant.

The T-cell clone designated F1.A.2 can recognize the moth cytochrome *c* synthetic fragment 86–89; 93–103(93E) (moth fragment) in association with either $E_{\beta}^k E_{\alpha}^k$ or $E_{\beta}^b E_{\alpha}^k$. These two Ia molecules differ by only four amino acids in the N-terminal domain of their β -chains (ref. 7 and P. Jones, personal communication). The moth fragment was 13-fold more potent (defined in Fig. 1) when recognized in association with $E_{\beta}^k E_{\alpha}^k$ than with $E_{\beta}^b E_{\alpha}^k$ (one representative experiment is shown in Fig. 2, an

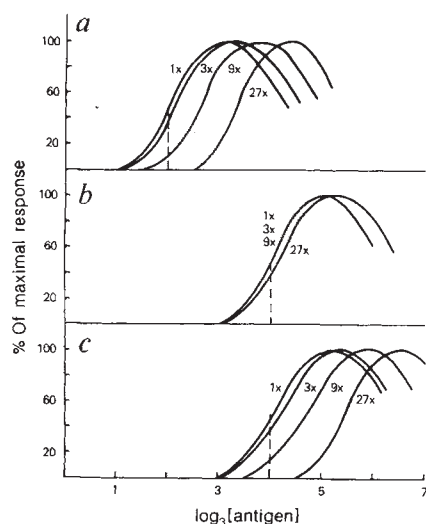


Fig. 1 Schematic representation of shifts in the antigen dose-response for various numbers of responding T cells. *a*, The proliferative response of a T-cell clone to varying concentrations of antigen in the presence of a fixed number of antigen-presenting cells (that is, the number of Ia molecules is constant). Each curve represents a different number of responding T cells, shown in threefold increments (1 $\times$, 3 $\times$, 9 $\times$, 27 $\times$). Because receptor-bearing cells compete with each other for ligand, as the number of responding cells in culture is increased, the fractional response will decrease³⁻⁶. When this results in an observable shift in the dose-response curves, the 'transition point' is said to have been reached³. In *a*, by inspection, the transition point is found to occur between the 1 $\times$ and 3 $\times$ increments. *b*, Illustration of a potency difference due to avidity. In the example shown, an antigen analogue exhibits a ninefold lower potency than the parent molecule. The relative potency of different antigen analogues (indicated here by a broken line perpendicular to the abscissa) is determined by comparing the concentrations of antigen required to achieve 50% of the maximal response for a specific responding cell number at or below the transition point. An alteration in the antigen that reduces the affinity of the antigen-specific receptor for antigen-Ia will result in a loss of potency. Furthermore, the lower affinity will also result in a decrease in the ability of the antigen-specific receptors to compete for ligand. Therefore, the number of responding cells required to achieve the transition point will increase (*b*). Alternatively, other modifications to the antigen might not affect the affinity of the antigen-specific receptor for the antigen-Ia, but would still cause a loss of potency (for example, by altering the way in which the antigen is processed, its solubility in the lipid membrane of the APC, its ability to interact with the Ia molecule) as shown in *c*. *c*, Dose-response curves using such an antigen analogue that differs in potency by ninefold from the analogue in *a*. In this case there is no difference in the affinity of the antigen-specific receptor for either the analogue or the original molecule. Therefore, the number of responding T cells needed to achieve the transition point is identical. The two different patterns shown here could also be generated by altering the affinity of the Ia molecule for either the receptor (*b*) or the antigen (*c*). Note that the descending portion of each of these schematic dose-response curves represents the phenomenon of 'high-dose suppression' seen on antigen activation of T cells¹⁶. This is not the result of antigen toxicity, but rather of receptor occupancy, and the shifts in the descending portions of the dose-response curves roughly parallel the shifts in the ascending portions.

average of four experiments is given in Table 1). We investigated the way in which varying the responding number of this T-cell clone affected its response to antigen in association with either Ia molecule. Despite the difference in potency, we found no significant difference in the degree to which the antigen dose-response curves shifted when antigen-presenting cells bearing either Ia molecule were used (Table 1). The unlikely possibility that the number of $E_\beta E_\alpha$ molecules present on the two types of antigen-presenting cells differed by 13-fold, thus accounting

Table 1 Summary of shifts in the antigen dose-response curve using moth 86-89; 93-103(93E) in association with $E_\beta E_\alpha^k$ or $E_\beta E_\alpha^k$

APC	Expt	[Ag] _{50%} ratio No. of responding T cells ($\times 10^{-3}$):			
		2	6	18	54
B10.A ($E_\beta E_\alpha^k$)	1	1 [0.00036]*	1.56	2.19	5.0
	2	1 [0.00018]	1.11	2.22	3.11
	3	1 [0.00014]	1.0	1.14	2.14
	4	1 [0.00075]	1.19	1.77	3.16
Geometric mean:		1 [0.00029]†	1.20 (1.10)	1.77 (1.17)‡	3.20 (1.19)‡
B10.A(5R) ($E_\beta E_\alpha^k$)	1	1 [0.0056]	1.27	2.32	5.36
	2	1 [0.0018]	1.39	3.11	3.94
	3	1 [0.0018]	1.0	1.39	2.50
	4	1 [0.011]	1.06	1.59	2.38
Geometric mean:		1 [0.0038]†	1.17 (1.08)	2.00 (1.20)‡	3.35 (1.21)‡

T-cell proliferation assays were performed as described in Fig. 2 legend. The concentration of antigen required to achieve 50% of the maximal stimulation of the clone F1.A.2 ([Ag]_{50%}) in the presence of either B10.A antigen-presenting cells (APC) (bearing the Ia molecule $E_\beta E_\alpha^k$) or B10.A(5R) APC (bearing $E_\beta E_\alpha^k$) at varying numbers of responding cells was determined in four separate experiments. A ratio was then calculated for each group by comparing the concentration of antigen required for a 50% response with the concentration required for a 50% response by the lowest number of responding T cells, 2×10^3 per well. Values in parentheses are $\times / \div$ s.e.m. of the geometric means.

* The concentration of antigen (μ M) required to achieve 50% of the maximal response at 2×10^3 responding T cells.

† The ratio of these two numbers ($0.0038/0.00029 = 13.1$) defines the potency difference for the two different antigen-Ia molecule ligands.

‡ Significantly different from 1 at the 5% level (two-tailed Student's *t*-test).

for the difference in potency, was ruled out by quantitative flow microcytofluorimetric analysis using a monoclonal antibody (14.4.4) that binds to the E_α chain; the levels of E_α expressed on the spleen cells from the two strains were comparable (data not shown).

A potency difference that results from an alteration in the affinity of the T-cell receptor for any component of its ligand (either the antigen or the Ia molecule) would be seen as a difference in avidity of the T-cell receptor for the entire ligand, and should be reflected in the antigen dose-response curve shifts that occur as the number of responding T cells is increased. As the ligand recognized with lower avidity must be present in greater amounts than the ligand recognized with higher avidity in order to achieve the same degree of activation, the competition that occurs between receptor-bearing cells should result in a greater fractional depletion of the ligand recognized with higher avidity. In the case where the Ia molecule has been altered, the shifts of the antigen dose-response curve would occur at lower numbers of responding T cells for the Ia molecule requiring less antigen than for the Ia molecule requiring more antigen. In contrast, if the functional difference between the two Ia molecules were a consequence of their interaction with the antigen, no difference would be expected between them in the way the antigen dose-response curve shifted. In this case, although the amount of soluble antigen would vary at the same degree of cellular response, the amount of the true ligand (antigen-Ia) would be the same. For the antigen-Ia molecule combination studied here (moth fragment plus $E_\beta E_\alpha^k$ and moth fragment plus $E_\beta E_\alpha^k$), no difference was seen in the shifts of the antigen dose-response curve, despite a 13-fold difference in the observed potency. This result suggests that there is no avidity difference between the T-cell antigen-specific receptor and these two ligands. The most likely mechanism by which a change in the molecule could cause the potency difference is through a change in the interaction between the Ia molecule and the antigen. Although altering the Ia molecule did not result in a change in avidity in this particular case, the possibility exists that, with other T-cell clones, making such a change will alter a site on the Ia molecule that makes contact with the T-cell receptor, and hence alter the affinity of the antigen-specific receptor for the Ia molecule.

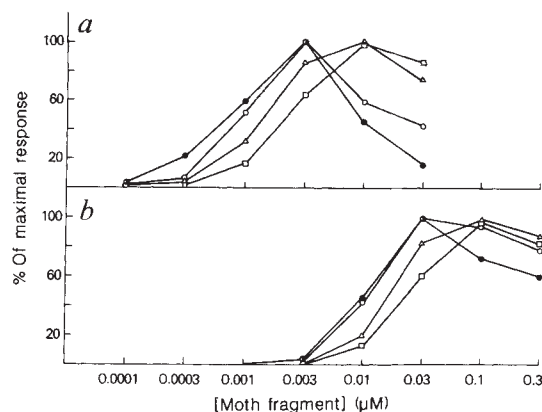


Fig. 2 The proliferative response of varying numbers of a T-cell clone to the antigen moth cytochrome *c* in the presence of the $E^k E^k_a$ Ia molecule (B10.A APC; *a*) or the $E^b E^k_a$ Ia molecule (B10.A(5R) APC; *b*). The T-cell clone F1.A.2 was obtained by Dr Gen Suzuki of our laboratory from [B10.A × B10.A(3R)] F_1 mice immunized with pigeon cytochrome *c* in Freund's complete adjuvant (Difco). Cell lines were established from T cells taken from the draining lymph nodes and T-cell clones prepared by limiting dilution as described previously³. The clone F1.A.2 was found to respond to a variety of cytochrome *c* species variants in association with either $E^k E^k_a$ or $E^b E^k_a$, as determined by its proliferative response to irradiated (3,300 rad) spleen cells from congenic C57BL/10 mice in the presence or absence of anti-Ia monoclonal antibodies specific for products of the *I-A* or *I-E* subregion (data not shown). The nomenclature used to describe the Ia molecules in this study is as follows: the E denotes the subregion of the mouse *H-2* complex that encodes the protein, the subscripts refer to the two chains that make up the Ia heterodimer, and the superscripts refer to the allelic forms of the two chains. The T-cell clone was maintained by periodic stimulation with pigeon cytochrome *c* plus irradiated B10.A spleen cells (for 4 days), followed by a period of rest (9–17 days) in the absence of antigen. Concanavalin A-stimulated rat spleen cell supernatant (as a source of interleukin-2) was occasionally added to the cultures 48 h after stimulation with antigen to promote optimal expansion of the cells. After a rest period, F1.A.2 T cells were titred into 96-well flat-bottomed plates (3596, Costar) at 2×10^3 (●), 6×10^3 (○), 18×10^3 (Δ) or 54×10^3 (□) cells per well. Each well contained 5×10^5 irradiated (3,300 rad) spleen cells from B10.A (*a*) or B10.A(5R) (*b*) mice as a source of APC bearing the Ia molecules, and varying concentrations of the synthetic peptide moth cytochrome *c* 86–89; 93–103(93E)¹⁷. Each well was brought to a final volume of 200 μ l with complete medium, a 50:50 (v/v) mixture of RPMI 1640 medium (Biofluids, Rockville, Maryland) and Eagle's Hank's Amino Acids (Biofluids), supplemented with 10% fetal calf serum, 5×10^{-5} M 2-mercaptoethanol, 4 mM glutamine, 100 U ml⁻¹ of penicillin and 150 μ g ml⁻¹ gentamicin. After 48 h of culture at 37°C in a humidified atmosphere containing 5% CO₂, the wells were pulsed with 1 μ Ci of ³H-thymidine (6.7 Ci mmol⁻¹; ICN); 16 h later the plates were harvested (PHD Cell Harvester, Cambridge Technologies, Cambridge, Massachusetts) and the amount of thymidine incorporation assessed by liquid scintillation counting. The results of duplicate cultures were averaged and plotted as a percentage of the maximal response achieved for each responding cell number. The maximum responses in this experiment were as follows. B10.A APC: 2×10^3 , 19,200 c.p.m.; 6×10^3 , 55,400 c.p.m.; 18×10^3 , 126,700 c.p.m.; 54×10^3 , 350,300 c.p.m. B10.A(5R) APC: 2×10^3 , 23,000 c.p.m.; 6×10^3 , 63,400 c.p.m.; 18×10^3 , 157,100 c.p.m.; 54×10^3 , 263,300 c.p.m.

The data presented here can be used to discriminate between some of the models proposed to explain the molecular basis of T-lymphocyte co-recognition. One model hypothesizes that the three molecular elements physically interact to form a stable complex that ultimately leads to signal transduction (the trimolecular complex model⁸; Fig. 3*a*). In this model, the T-cell antigen-specific receptor, along with the antigen and the Ia molecule, combine to deliver a stimulatory signal to the T cell. A change in affinity at any one of the potential interaction sites

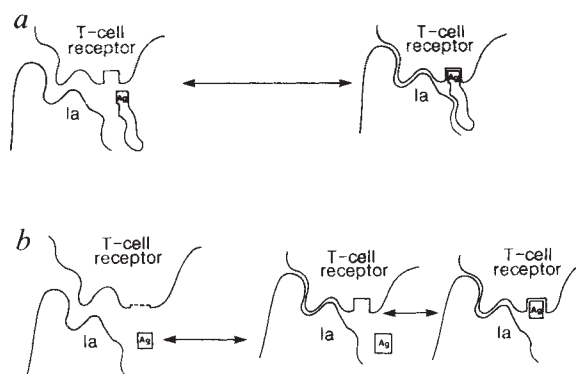


Fig. 3 Two models for antigen-specific T-cell receptor recognition of antigen and an Ia molecule. *a*, The trimolecular complex model. The three elements of the trimolecular complex (antigen-specific receptor, antigen and Ia molecule) are in equilibrium with a ternary complex. There are three potential sites of interaction: T-cell receptor/antigen, T-cell receptor/Ia molecule, and antigen/Ia molecule⁶. Each bimolecular interaction in the complex can be considered as a discrete reaction with its own kinetic constants. *b*, The allosteric or induced-fit model. In an allosteric model, the T-cell antigen-specific receptor has two binding sites: one for the antigen and the other for the Ia molecule. In an induced-fit model, a binding site exists for only one of the elements in the absence of the other. In either case, the affinity for one of these elements (the antigen, in this example) is so low that at equilibrium, insufficient binding occurs to lead to activation. The antigen-binding portion of the T-cell receptor is depicted by a shallow dotted depression to indicate that the antigen-binding site might not exist before the binding of the Ia molecule, or that it might exist predominantly in a low-affinity state (left). In this example, when the Ia molecule occupies its portion of the receptor, a conformational change occurs (or an alternative state is stabilized) (middle) such that the receptor can now bind the antigen with high affinity (right). This results in signal transduction. There is no postulated direct contact between the antigen and the Ia molecule. Note that a completely allosteric model requires two alternative states after Ia molecule binding to account for the differences we have observed previously in the fine specificity of antigen binding in the presence of $E_a^b E_\alpha^k$ as opposed to $E_a^b E_\alpha^k$ Ia molecules.¹⁸

would alter the potency of the ligand by virtue of the effect on the stability of the final trimolecular complex. If only the interaction between the Ia molecule and the antigen were altered, the affinity of the T-cell antigen-specific receptor for its ligand (antigen-Ia) could be unchanged, and the antigen dose-response curves would shift identically regardless of which Ia molecule was used to stimulate the T cell; this was the result obtained for clone F1.A.2. In contrast, the present results are not consistent with a simple allosteric or induced-fit model (Fig. 3b) of T-cell recognition. In such a model, differences in the ability of one antigen to stimulate a T-cell clone in the presence of two different Ia molecules must result from a difference in the affinity of the T cell's antigen-specific receptor for either the two Ia molecules or the antigen (the latter due to different states of the receptor induced or stabilized by the two Ia molecules)^{9,10}.

The results presented here argue only that the Ia molecule and the antigen interact physically, thus they do not exclude a third model, that of 'altered self', in which the antigen induces a conformational change in the Ia molecule which is detected by the T cell's antigen-specific receptor^{11,12}. This particular model seems unlikely in view of the accumulating data which suggest that the antigen-specific receptor can bind antigen in the absence of an Ia molecule under certain circumstances¹³⁻¹⁵.

We thank Drs Ronald Germain and Barbara Fox for critical review of the manuscript, and Ms Shirley Starnes for preparation of the manuscript.

Received 15 October 1985; accepted 27 January 1986.

1. Schwartz, R. H. *Adv. Immun.* **38**, 31-201 (1986).
2. Babbitt, B. P., Allen, P. M., Matsueda, G., Haber, E. & Unanue, E. R. *Nature* **317**, 359-361 (1985).
3. Ashwell, J. D., Fox, B. S. & Schwartz, R. H. *J. Immun.* **136**, 757-768 (1986).
4. Cuatrecasas, P. & Hollenberg, M. D. *Adv. Protein Chem.* **30**, 251-451 (1976).
5. Moyle, W. E., Lee, E. Y., Bahl, O. P., Garfink, J. E. & Rodbard, D. *Am. J. Physiol.* **232**, E274-E285 (1977).
6. Moyle, W. R., Lee, E. Y., Bahl, O. P. & Rodbard, D. in *Receptors and Hormone Action* (Academic, New York, 1978).
7. Mengle-Gaw, L. & McDevitt, H. O. *A. Rev. Immun.* **3**, 367-396 (1985).
8. Heber-Katz, E., Hansburg, D. & Schwartz, R. H. *J. molec. cell. Immun.* **1**, 3-14 (1983).
9. Schwartz, R. H., Heber-Katz, E. & Hansburg, D. in *Intercellular Communication in Leukocyte Function* (Wiley, Chichester, 1983).
10. Cleveland, W. L. & Erlanger, B. F. *Molec. Immun.* **21**, 1037-1046 (1984).
11. Zinkernagel, R. M. & Doherty, P. *Nature* **251**, 547-548 (1974).
12. Sherman, L. *Nature* **292**, 511-513 (1982).
13. Carel, S., Bron, C. & Corradin, G. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4832-4836 (1983).
14. Rao, A., Weng-Ping, W., Faes, S. J. & Cantor, H. *Cell* **36**, 879-888 (1984).
15. Siliciano, R. F., Keegan, A. D., Dintzis, R. Z., Dintzis, H. M. & Shin, H. S. *J. Immun.* **135**, 906-914 (1985).
16. Matis, L. A., Glimcher, L. H., Paul, W. E. & Schwartz, R. H. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6019-6023 (1984).
17. Schwartz, R. H., Fox, B. S., Fraga, E., Chen, C. & Singh, B. *J. Immun.* **135**, 2498-2608 (1985).
18. Schwartz, R. H. *A. Rev. Immun.* **3**, 237-261 (1985).

T-cell-mediated association of peptide antigen and major histocompatibility complex protein detected by energy transfer in an evanescent wave-field

Tania H. Watts, Hermann E. Gaub*
& Harden M. McConnell

Stauffer laboratory for Physical Chemistry, Stanford University,
Stanford, California 94305, USA

Helper T cells recognize foreign antigen displayed on antigen-presenting cells which also express self-molecules of the major histocompatibility complex (MHC)¹. A single T-cell receptor mediates recognition of both MHC and foreign antigen^{2,3}. A proposed ternary complex between T-cell receptor, foreign antigen and MHC antigen has not yet been demonstrated (see ref. 1 for review). Here, we show that a fluorescein-labelled synthetic peptide, together with Texas red-labelled class II MHC antigen, I-A^d, stimulates the production of interleukin-2 by a peptide-specific I-A^d-restricted T-cell hybridoma when reconstituted in a lipid membrane on a glass substrate. Under the same conditions, resonance-energy transfer from donor peptide to acceptor I-A can be stimulated in an evanescent wave-field only in the presence of the specific T-hybrid. Our results show that the T cell stabilizes an association between peptide antigen and class II MHC protein to within a distance of about 40 Å.

The T-cell hybridoma 3DO-54.8 recognizes a 17-residue tryptic peptide representing residues 323-339 of ovalbumin when processed ovalbumin is presented by I-A^d-bearing antigen-presenting cells⁴. Previous work has shown that when phospholipid vesicles are allowed to interact with a clean glass slide, the vesicles fuse to form a continuous membrane^{5,6}. When these vesicles contain MHC antigens, the resulting planar membranes are effective in stimulating T-cell responses^{5,6}. When synthetic peptide 323-339 is presented by I-A^d-containing planar membranes, the midpoint of the dose-response curve occurs at 2 µM antigen⁷. The N-terminal three residues and the C-terminal three residues of this peptide can be removed with only a small shift in the dose-response curve⁷. In the present study, the amino-terminus of peptide 323-339 was modified with fluorescein 5-isothiocyanate (FITC). The resulting peptide was indistinguishable from unlabelled peptide in stimulating interleukin-2 (IL-2) production from 3DO-54.8 cells (data not shown). The I-A^d was

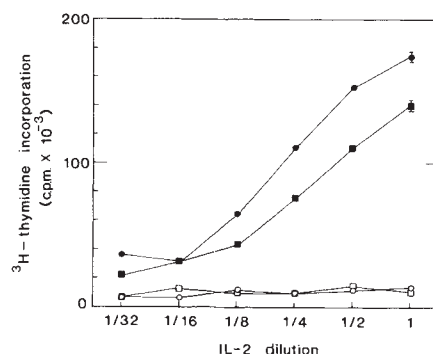


Fig. 1 T-cell stimulation by Texas red-labelled I-A^d in supported lipid membranes. IL-2 production from 3DO-54.8 cells in contact with supported planar membranes containing purified I-A^d was measured using the IL-2-dependent cell line CTLL as described previously⁶. Planar membranes contained unlabelled I-A^d (●, ○) or Texas red-labelled I-A^d (five Texas red molecules per I-A molecule) (■, □). Filled symbols, fluorescein-labelled peptide antigen added to a final concentration of 2 µM; open symbols, no antigen. Peptide labelling was carried out for 2 h at room temperature in 0.1 M sodium bicarbonate/carbonate buffer pH 9.0 using a fivefold molar excess of FITC (isomer 1, Molecular Probes Inc., Junction City) over peptide. Labelled peptide was purified using HPLC⁷. Texas red (Molecular Probes Inc.) labelling of I-A was carried out for 2.5 h on ice in 0.5% (w/v) sodium deoxycholate in 0.1 M sodium bicarbonate/carbonate pH 9.5 using a 1,000-fold molar excess of Texas red to I-A. Excess Texas red was removed by exhaustive dialysis. The Texas red-labelled I-A was reconstituted with egg phosphatidylcholine/cholesterol as described previously⁶. Dialysis of a lipid sample containing free Texas red showed that all the noncovalently linked Texas red could be removed by the exhaustive dialysis.

modified with the fluorescent probe Texas red⁸. Usually, the product contained five Texas red molecules per I-A molecule. Figure 1 shows IL-2 release by 3DO-54.8 cells in contact with planar membranes containing purified unlabelled or Texas red-labelled I-A. The results shown are for 2 µM fluorescent peptide 323-339. Although there is a decrease in the stimulatory ability of I-A after modification with Texas red, the overall response is still high.

I-A was incorporated into planar membranes on two classes of glass (or quartz) substrates. The experiments described in Fig. 1 were carried out on hydrophilic supports as described previously^{6,7}. In addition, planar membranes were prepared by fusing I-A^d-containing phospholipid vesicles on supports that had been alkylated with hexadecyltrichlorosilane (C₁₆ supports). (Alkylated supports have been described elsewhere⁹). The resulting membranes contain I-A^d as measured by Texas red fluorescence. IL-2 production by 3DO-54.8 cells in contact with planar membranes on C₁₆ supports requires ~10-fold more peptide antigen than the same membranes on hydrophilic supports (data not shown). A preliminary experiment indicates that the reconstituted I-A on C₁₆ supports exhibits slow lateral diffusion (10⁻¹⁰ cm² s⁻¹) whereas lateral diffusion of MHC antigens in lipid bilayers on hydrophilic supports is below the limits of detection (<10⁻¹¹ cm² s⁻¹)^{5,6}. These values were determined by using fluorescein-labelled anti-I-A antibody and measuring fluorescence recovery after photobleaching¹⁰.

The apparatus for observing resonance energy transfer at the T-cell/lipid membrane interface is shown in Fig. 2a. Essential features are the use of evanescent wave excitation^{11,12} and the use of a monochromator to record fluorescence emission spectra. Energy transfer between donor peptide and acceptor MHC molecules will be stimulated only if these molecules are within 40 Å of one another (for a review of Förster energy transfer see refs 13, 14). Only those donor-labelled peptide molecules that are within 800 Å of the planar membrane surface are excited; this has the effect of reducing the fluorescein background in our samples by three orders of magnitude and also greatly reduces

* Present address: Institute for Biophysics E22, Department of Physics, Technische Universität München, 8046 Garching, FRG.

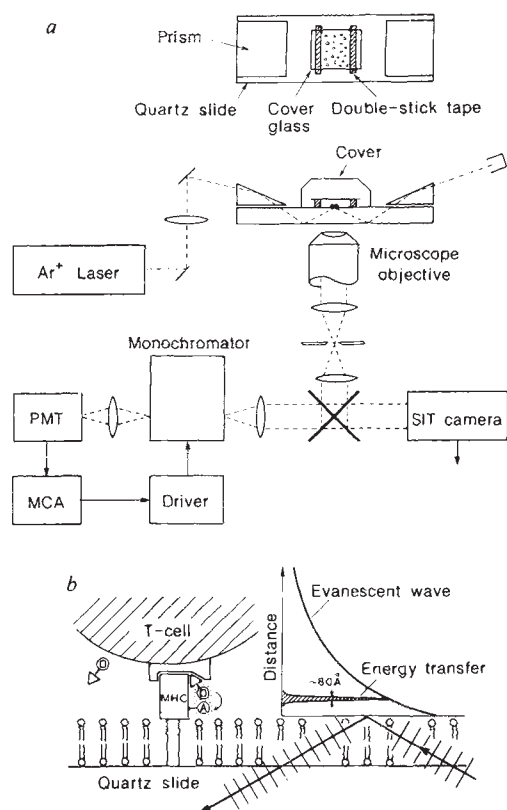


Fig. 2 *a*, Experimental design; *b*, schematic representation of the molecular problem. *a*, The microfluorometer is based on an inverted microscope (Nikon Diaphot TMD) on an optical isolation table. The beam of an argon ion laser (Spectra Physics Ar 2000) is attenuated to <10 mW by a set of partially reflecting mirrors and a half-wave plate followed by a Glan-Thompson prism. The *p*-polarized beam is focused ($f=600$ mm) and coupled by a high-refractive-index prism ($n=1.8$) into a quartz microscope slide. The prism is positioned so that the first reflection on the upper side of the slide is in the sample area. Fine adjustment of the mirrors positions the excitation spot (~ 200 μ m diameter) in the field of view of the microscope. A $\times 40$ variable correction 0.55 NA Nikon objective was used for all measurements. A variable diaphragm in the image plane limits the area from which fluorescence is detected to a spot of $<150 \times 150$ μ m. The image of this diaphragm is focused ($f=600$ mm) with the same aperture angle as that on the entrance slit of the monochromator (Spex 1681). The output slit is enlarged threefold and projected on the cathode of a photomultiplier (RCA 31034 cooled housing, -30 $^{\circ}$ C). Photons are counted with a multi-channel analyser (MCA; LeCroy 3500) with an EG&G 436 discriminator and an EG&G 535 preamplifier module. The monochromator is driven by a Spex Compudrive. PMT, photomultiplier tube; SIT silicon-intensified target. A plane wave hitting the boundary to an optically less dense medium is totally internally reflected if the incidence angle exceeds the critical angle (see ref. 12 for review). Lipid and the quartz have a very similar refractive index and are treated as one medium. The reflected wave creates an exponentially decaying evanescent electrical field in the optically less dense medium, the $1/e$ distance is ~ 800 Å. Only those molecules that are close to the interface are excited. The acceptor fluorophore (designated A) is attached to the MHC molecule, confined to the planar membrane, as indicated. The donor fluorophore (D) is attached to the peptide antigen. If the T-cell receptor brings about a significant enhancement in the probability that the donor and acceptor groups are within 40 Å of each other, then energy extracted from the evanescent wave field and absorbed by the donor group D leads to an enhancement in the fluorescence of A through resonance energy transfer. An enhancement of emission from A does not distinguish between a model for a ternary complex in which there are two closely apposed, but separate, sites for antigen and MHC on the T-cell receptor, from one in which there is one continuous site for binding the two molecules.

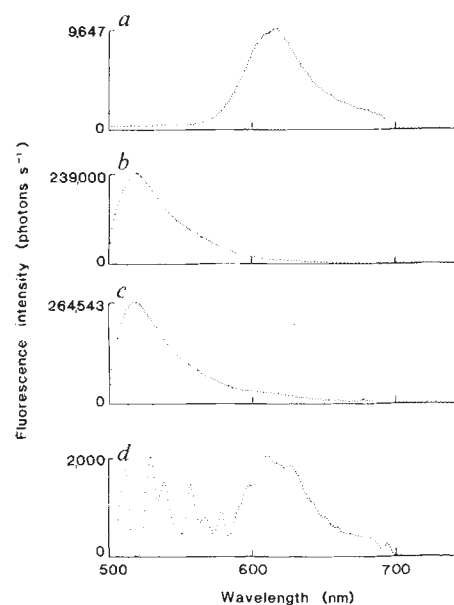


Fig. 3 Energy transfer from fluorescein-labelled peptide to Texas red-labelled I-A^d in the presence of 3DO-54.8 cells. Shown are typical raw data obtained from the apparatus described in Fig. 2. *a*, Spectrum of a planar membrane containing Texas red-I-A^d and lipid alone; *b*, reference spectrum of fluorescein-labelled peptide antigen and 3DO-54.8 cells in contact with unlabelled I-A^d-containing planar membranes; *c*, 3DO-54.8 cells in contact with the planar membrane shown in *a* in the presence of 20 μ M fluorescein-labelled peptide; *d*, difference spectrum $c-b-a$. Spectra *b* and *c* were scaled before subtraction using the integrals of the curves from 500 to 544 nm. The Texas red signal is negligible at these wavelengths. From this difference spectrum the Texas red-I-A signal was then subtracted directly. The noise in the difference spectrum in the wavelength region 500–580 nm arises from the subtraction of two large numbers, which can be caused by 1% fluctuations in laser intensity. In the region of the Texas red emission, this noise is less because the overall intensities used in the subtraction are 10-fold smaller. The fact that the double difference spectrum has the characteristic line shape of the Texas red emission shows the fidelity of the subtraction procedure. The double difference spectrum in the region from 500 to 580 nm is random noise. Cells and peptide were suspended in buffer consisting of 10 mM HEPES, 120 mM NaCl, 5.4 mM KCl, 5.6 mM glucose, 2.0 mM CaCl₂, 1.5 mM MgCl₂, 0.5–2% FCS. Cells were allowed to settle for 7 min at room temperature before the spectrum was recorded.

the problem of cell autofluorescence. From the overlap integral between the emission of fluorescein and the absorption of Texas red, the energy-transfer distance is calculated to be 40 Å (ref. 13).

As a positive control, energy transfer from fluorescein-labelled anti-I-A antibodies was examined. When specific anti-I-A^d antibody (MKD6)² was used, a 90% enhancement of the Texas red signal was observed with antibody labelled with 8 fluorescein molecules per antibody. A 25% enhancement of the Texas red signal was observed when the antibody was labelled at a ratio of 1:1. No enhancement of Texas red fluorescence was observed with nonspecific fluorescein-labelled antibody (anti-I-A^k antibody 10-12.6 (ref. 15) labelled with 5 fluorescein molecules per protein) (data not shown).

Texas red/I-A-containing planar membranes were prepared as described previously⁶ and their fluorescence emission spectra recorded (Fig. 3*a*). Without moving the sample on the microscope stage, cells and peptide were added. Fluorescein-labelled peptide antigen was pre-mixed with T cells to give a final concentration of 20 μ M peptide, 0.5–2% fetal calf serum (FCS) and sufficient cells to produce a confluent layer. In some experiments anti-I-A antibody or a 10-fold excess of unlabelled peptide was included in the solution (see Fig. 3*c*). The spectrum of a control sample consisting of fluorescein-labelled peptide

Table 1 Summary of energy transfer results

Expt	No. of trials	Support*	Concentration FITC-peptide (μM)†	Concentration unlabelled peptide (μM)‡	Antibody§	Cells	Increase in Texas red emission $\pm$ s.d.¶
a	4	H, C ₁₆	20	—	—	—	0 $\pm$ 0
b	4	H	1	—	—	3DO-54.8	-13 $\pm$ 15
c	10	H	20	—	—	3DO-54.8	18 $\pm$ 8.5
d	3	H	20	200	—	3DO-54.8	4 $\pm$ 6
e	10	C ₁₆	20	—	—	3DO-54.8	28 $\pm$ 10
f	4	C ₁₆	20	200	—	3DO-54.8	-0.8 $\pm$ 15
g	2	C ₁₆ , H	20	—	MKD6	3DO-54.8	-20 $\pm$ 28
h	3	C ₁₆ , H	20	—	—	FS11-72.3	6 $\pm$ 10
i	1	H	20	—	—	CTLL	0

* H, hydrophilic support, resulting in a supported lipid bilayer; C₁₆, a support alkylated with hexadecyltrichlorosilane, resulting in a supported lipid monolayer.

† FITC-peptide, fluorescein-labelled synthetic peptide representing residues 323-339 of ovalbumin, modified and purified as described in Fig. 1. ‡ Unlabelled peptide, the same peptide without the fluorescein tag. The competing unlabelled peptide was added in 10-fold molar excess to the mixture of T cells and fluorescent peptide before adding them to the planar membrane.

§ Sufficient MKD6 antibody was added to the mixture of T cells and peptide to saturate the I-A with antibody.

|| 3DO-54.8, specific T cell (see text); FS11-72.3, self-reactive T-cell hybridoma which recognizes I-A^k; CTLL is the cytotoxic IL-2 indicator line.

¶ The per cent enhancement of the Texas red signal was determined from a comparison of the double-difference spectrum described in Fig. 3d with the original Texas red spectra.

plus 3DO-54.8 cells in contact with a planar membrane containing unlabelled I-A was recorded in a separate experiment (Fig. 3b). To subtract the fluorescein component from spectrum *c* to obtain the signal resulting from Texas red alone, spectra *b* and *c* were scaled using the integral of the curve between 500 and 544 nm. The signal obtained in spectrum *a* (obtained before addition of cells and peptide) was subtracted from the difference spectrum (*c* - *b*) to give the double difference spectrum *d*. This spectrum (*d* = *c* - *b* - *a*) is the enhancement of the Texas red signal caused by the addition of cells and peptide.

From the experiments summarized in Table 1, we conclude that: (1) In the absence of cells, there is no observed energy transfer between peptide and I-A at the same concentrations used in the experiments with cells (expt *a*), irrespective of whether the support is C₁₆ or hydrophilic. (2) At peptide concentrations representing the midpoint of the dose-response curve for IL-2 release, there is no detectable energy transfer in the presence of cells (expt *b*). (3) We always observe an increase in Texas red emission when peptide is added to a concentration of 20 μM together with the specific T cells. The magnitude of this enhancement is 10-30% on hydrophilic supports and 20-50% on C₁₆ supports (expt *c*). (4) When the energy transfer experiment is performed in the presence of a 10-fold molar excess of unlabelled peptide, we always observe a smaller enhancement of the Texas red emission signal (expts *d*, *f*). Each competition experiment was carried out immediately following a positive experiment and always showed a decrease in the effect, excluding the possibility that the enhancement of Texas red fluorescence is caused by a change in its quantum yield on complexation. (5) Anti-I-A^d antibody abolishes T-cell-mediated energy transfer (expt *g*). (6) Other T cells, not specific for I-A^d and ovalbumin, show little or no energy transfer (expts *h*, *i*). When a small amount of transfer was observed, this could not be diminished by addition of unlabelled peptide (data not shown). Recent experiments indicate that nonspecific trapping of the peptide between the two membranes leading to energy transfer, can be decreased by increasing the concentration of FCS. The only exception to the above statements was obtained with the T cell 3DO-18.3. This T cell comes from the same fusion as 3DO-54.8, responds to I-A^d and ovalbumin, but does not respond to the 17-residue synthetic peptide¹⁶. This cell mediates some energy transfer between I-A and peptide and there was competition with excess unlabelled peptide.

Our results provide compelling evidence that a specific T-helper cell stabilizes the proximity of peptide antigen and I-A^d incorporated in a lipid membrane. This result is particularly

significant in that it was obtained under conditions where a physiological response can be demonstrated. The results support theoretical models (see ref. 1 for review) of MHC-restricted recognition of antigen involving a T-cell receptor-mediated association of antigen and class II MHC molecules to within distances of the order of 40 Å. Without a clonotypic antibody against 3DO-54.8 cells we cannot say that the T-cell receptor α/β heterodimer (see ref. 17 for review) is responsible for this effect. In the absence of cells, no interaction between antigen and MHC was detected. Either there is no complex between antigen and MHC in the absence of the specific T cell or the affinity is too low to measure using this approach for this particular antigen/MHC combination.

At present we cannot distinguish whether the observed energy transfer arises from many antigen-MHC associations at relatively large distances (40 Å) where energy transfer is inefficient or whether there are a few associations where the energy transfer is efficient. We anticipate that these and other questions can be answered by extension of the techniques described here.

We thank J. Gariépy and G. Schoolnik for peptide synthesis and P. Marrack and J. Kappler for T cells. This work was supported by NIH grant 5RO1 A113587, Department of Defense equipment grant N00014-84-G-0210 and San Francisco Laser Center grant UC-SC-32072(H.M.M.); the MRC of Canada and the Alberta Heritage Foundation for Medical Research (T.H.W.); and by the Deutsche Forschungsgemeinschaft (H.E.G.).

Received 28 October 1985; accepted 10 January 1986.

- Schwartz, R. A. *Rev. Immun.* **3**, 237-261 (1985).
- Kappler, J. W., Skidmore, B., White, J. & Marrack, P. *J. exp. Med.* **153**, 1198-1214 (1981).
- Yagüe, J. *et al. Cell* **42**, 81-87 (1985).
- Shimonkevitz, R., Colon, S., Kappler, J. W., Marrack, P. & Grey, H. M. *J. immun.* **133**, 2067-2074 (1984).
- Brian, A. A. & McConnell, H. M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6159-6163 (1984).
- Watts, T. H., Brian, A. A., Kappler, J. W., Marrack, P. & McConnell, H. M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7564-7568 (1984).
- Watts, T. H., Gariépy, J., Schoolnik, G. & McConnell, H. M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5480-5484 (1985).
- Titus, J., Haugland, R., Sharrow, S. O. & Segal, D. M. *J. immun. Meth.* **50**, 193-204 (1982).
- Seul, M., Subramaniam, S. & McConnell, H. M. *J. phys. Chem.* **89**, 3592-3595 (1985).
- Smith, L. M., Rubenstein, J. L. R., Parce, J. W. & McConnell, H. M. *Biochemistry* **19**, 5907-5911 (1980).
- Weis, R. M., Balakrishnan, K., Smith, B. A. & McConnell, H. M. *J. biol. Chem.* **257**, 6440-6445 (1982).
- Axelrod, D., Burghardt, T. P. & Thompson, N. L. *A. Rev. Biophys. Bioengng* **13**, 247-268 (1984).
- Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 303-339 (Plenum, New York, 1985).
- Stryer, L. *A. Rev. Biochem.* **47**, 819-846 (1978).
- Oi, V. T., Jones, P. P., Goding, J. W., Herzenberg, L. A. & Herzenberg, L. A. *Curr. Topics Microbiol. Immun.* **81**, 115-129 (1978).
- Shimonkevitz, R., Kappler, J., Marrack, P. & Grey, H. *J. exp. Med.* **158**, 303-316 (1983).
- Haskins, K., Kappler, J. & Marrack, P. *A. Rev. Immun.* **2**, 51-66 (1984).

Specific interaction between the p53 cellular tumour antigen and major heat shock proteins

Orit Pinhasi-Kimhi*, Dan Michalovitz*,
Avri Ben-Zeev† & Moshe Oren*

Departments of Chemical * Immunology and † Genetics,
Weizmann Institute of Science, Rehovot 76100, Israel

The protein p53 is capable of participating in neoplastic transformation¹⁻³ and can form specific complexes with the large-T antigen of simian virus 40 (SV40)⁴⁻⁶. This interaction probably results in the stabilization of p53 (refs 7, 8) and may contribute to SV40-mediated transformation^{9,10}. Several non-SV40-transformed cells also exhibit a stabilized p53 which is present in elevated levels¹¹⁻¹³. Recently, this stabilization was shown to coincide with the ability to precipitate a polypeptide (p68) of relative molecular mass (M_r) 68,000–70,000 by anti-p53 monoclonal antibodies¹³⁻¹⁵. We now report that this co-precipitation indeed represents a specific complex between the two proteins; the complex sediments on a sucrose gradient as a relatively broad peak of 10–14S and can be dissociated *in vitro*. Furthermore, p68 is the HSP70 heat shock protein cognate, found in elevated levels in a p53-overproducing cell line. On heat-shock treatment of such overproducers, p53 also forms a complex with the related highly inducible HSP68.

The present studies used clone 6, a cell line derived by co-transformation of rat embryo fibroblasts with murine p53 and activated Ha-ras (D. Eliyahu *et al.*, in preparation). In such cells, p53 is both overproduced and stabilized (D. Eliyahu *et al.*, in preparation). Immunoprecipitation of labelled clone 6 proteins with anti-p53 monoclonal antibodies (Fig. 1a, lane 2) produces a prominent p53 band, along with a co-precipitating band of $M_r \sim 70,000$, previously¹⁵ referred to as p68. The precipitation of p68 could be explained in one of the following ways: (1) p68 is an irrelevant protein accidentally sharing an epitope with p53; (2) p68 is a modified more slowly migrating form of p53; (3) p68 forms a specific physical association with p53. In the first case one would not expect p68 to react with two different anti-p53 monoclonal antibodies. However, p68 is brought down by both PAb421¹⁶ and RA3-2C2¹⁷ antibodies, which bind to different domains of the p53 molecule^{18,19} (Fig. 1a, lanes 4, 5).

To elucidate the relationship between p53 and p68, a labelled clone 6 extract was denatured before incubation with anti-p53 monoclonal antibodies, to dissociate non-covalent complexes. This totally abolished the immunoprecipitation of p68 while having no effect on the reactivity of p53 (Fig. 1a, lane 3), as expected if p68 forms a physical complex with p53 rather than reacting directly with the monoclonal antibodies. Finally, comparison of the partial proteolytic patterns of p53 and p68²⁰ showed the patterns to be totally different, demonstrating that these are not closely related polypeptides (data not shown). Hence, the presence of p68 in the immunoprecipitate can be accounted for only if it is in physical association with p53, forming part of an anti-p53 reactive complex.

Figure 1a shows that there seems to be much more p53 than p68 in the precipitate, suggesting that only a minor fraction of p53 is tightly associated with p68. However, when immunoprecipitated proteins were visualized by Coomassie-blue staining (Fig. 1b, lane 1), large quantities of p68 were indeed seen, consistent with the notion that most of the p53 molecules are involved in the interaction.

The finding of a p53–p68 complex raised the possibility that the latter protein is related to the transforming activity of p53. It was therefore of interest to characterize p68, and for this, an immunoprecipitate of clone 6 cells was subjected to two-dimensional electrophoresis. A preliminary experiment (not

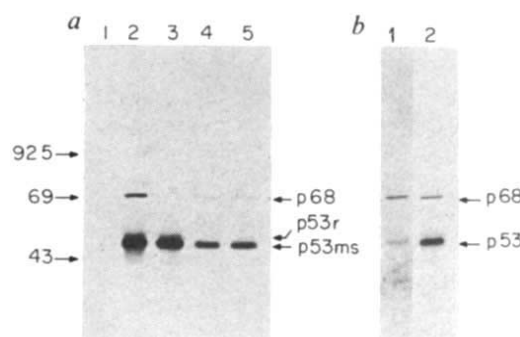


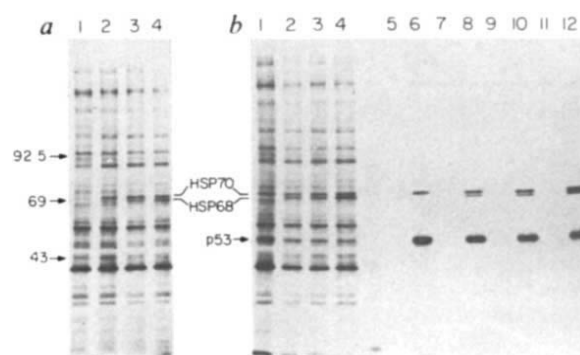
Fig. 1 Analysis of p53–p68 interaction. **a**, Immunoprecipitation with different antibodies. ³⁵S-labelled proteins were extracted from cells of clone 6, derived by transformation of rat embryo fibroblasts by p53 plus Ha-ras. Equal amounts of trichloroacetic acid (TCA)-insoluble radioactivity (1.4×10^6 c.p.m.) were reacted with either non-immune serum (lane 1) or anti-p53 monoclonal antibody PAb421 (lane 2). An identical sample was subjected to heat denaturation in reducing conditions³⁵ and similarly reacted with PAb421 (lane 3). The supernatant (unreacted material) of the immunoprecipitate shown in lane 2 was divided in half and incubated again with either the anti-p53 monoclonal antibody RA3-2C2 (lane 4) or PAb421 (lane 5). The autoradiograms are shown. Numbers on the left denote the sizes and positions of co-electrophoresed M_r markers. p53ms, mouse p53; p53r, endogenous rat p53. **b**, Comparison of stained and radiolabelled p53 and p68. A radiolabelled extract of clone 6, derived from one-quarter of a 90-mm dish, was reacted with PAb421 and the reacting polypeptides were resolved on a 12.5% SDS-polyacrylamide gel. The Coomassie blue-stained gel section (lane 1) and the autoradiogram of the same section (lane 2) are shown. Autoradiography was performed without prior fluorography.

Methods. Clone 6 is a cell line established from a transformed focus generated by co-transformation of rat embryo fibroblasts with activated Ha-ras and the p53-specific plasmid pLTRp53cG (refs 1, 26 and D. Eliyahu *et al.*, in preparation). For the experiment shown in **a**, a confluent 60-mm dish was labelled with 50 μ Ci ³⁵S-methionine for 3 h and extracts were prepared and analysed by immunoprecipitation followed by SDS-polyacrylamide gel analysis and autoradiography as previously described³⁹. The aliquot analysed in lane 3 was first made 5% in β -mercaptoethanol and 0.5% in SDS, heated to 100 °C for 5 min, cleared at 12,500g for 10 min and the supernatant was taken for immunoprecipitation. For experiment **b**, the material to be analysed was derived from one-quarter of a 90-mm dish, labelled with 50 μ Ci of ³⁵S-methionine for 3 h. Other details were as for **a**.

shown) revealed that the position of p68 coincided with that of an abundant cellular polypeptide, migrating similarly to the $\sim 70,000$ - M_r rodent heat shock proteins (HSPs). Furthermore, at least in one system, two polypeptides of $M_r \sim 68,000$ and $\sim 70,000$ co-precipitated with p53¹⁵, reminiscent of the HSP doublet reported in various mammalian species²¹⁻²⁴. To determine the relationship between p68 and HSPs, clone 6 cells were subjected to brief periods of hyperthermia²², followed by radiolabelling and protein analysis. A similar treatment was performed on Rat-1 established rat fibroblasts²⁵, which produce low amounts of p53²⁶. As Fig. 2 shows, p68 indeed co-migrated with a major HSP, the semi-inducible HSP70. Moreover, on increased hyperthermia, increasing amounts of a slightly faster-migrating polypeptide became prominent in the immunoprecipitate. This polypeptide co-migrated exactly with the highly inducible HSP68. Both polypeptides were absent from immunoprecipitates of denatured samples (data not shown), indicating that this second protein also complexed with p53. Neither of the two $\sim 70,000$ - M_r bands was convincingly detectable in immunoprecipitates of heat-shocked Rat-1 cells, although the existence of minute amounts of the complex with p53 could not be excluded (data not shown). To confirm the identity between p68 and HSP70, immunoprecipitates were resolved on two-dimensional gels and compared with total extracts from cells

Fig. 2 Analysis of p53 co-precipitating proteins following heat shock. Non-transformed Rat-1 cells (*a*) or clone 6 cells (*b*) were subjected to increasing temperatures followed by radiolabelling and extraction. Equal amounts of TCA-insoluble radioactivity (31,000 c.p.m.) were analysed directly by gel electrophoresis (lanes 1–4). Parallel but larger aliquots (7.2×10^5 c.p.m.) were reacted with either non-immune serum (lanes 5, 7, 9, 11) or anti-p53 monoclonal antibody PAb421 (lanes 6, 8, 10, 12). Hyperthermia was performed at the following temperatures: 37 °C (lanes 1, 5, 6), 44 °C (lanes 2, 7, 8), 45 °C (lanes 3, 9, 10) or 46 °C (lanes 4, 11, 12). The positions of the major rat heat-shock proteins HSP70 and HSP68 are indicated.

Methods. Subconfluent 60-mm dishes of either Rat-1²⁵ or clone 6 cells were pre-starved for 45 min in non-radioactive methionine-free medium. For the last 10 min of this period, the dishes were subjected to hyperthermia as described previously²². Incubation was then resumed at 37 °C for 4 h in methionine-free medium containing either 30 μ Ci (37–45 °C dishes) or 60 μ Ci (46 °C dish) of ³⁵S-methionine. Extracts were prepared as described previously and small aliquots analysed directly by SDS-polyacrylamide gel electrophoresis. Most of the extracts were used for immunoprecipitation and analysed as for Fig. 1.



either exposed to hyperthermia or kept at 37 °C (Fig. 3). The pertinent heat-shock proteins are most easily identified in total extracts from Rat-1 cells (Fig. 3e,f): HSP70 is easily detectable at 37 °C but is several-fold overproduced at 46 °C, whereas HSP68 is seen only after hyperthermia. On the other hand, clone 6 shows hardly any increase in HSP70 after heat shock (Fig. 2, Fig. 3c,d), while HSP68 displays the same pattern as in Rat-1. Most importantly, comparison of the total cell extracts with the immunoprecipitates reveals that p53 indeed co-migrates precisely with HSP70, while the other polypeptide corresponds to HSP68. This is shown most clearly in Fig. 3g, in which an immunoprecipitate was mixed and co-electrophoresed with a small amount of total cell extract, to allow a better alignment of the co-precipitating polypeptides with the HSPs.

To characterize the complex better, a clone 6 extract was subjected to sucrose gradient sedimentation (Fig. 4). While most cellular proteins were in monomeric form (fractions 20–23), there was hardly any monomeric p53. Rather, p53 was found in oligomers sedimenting as a relatively broad peak of 10–14S, with a small proportion of larger complexes. HSP70 was clearly

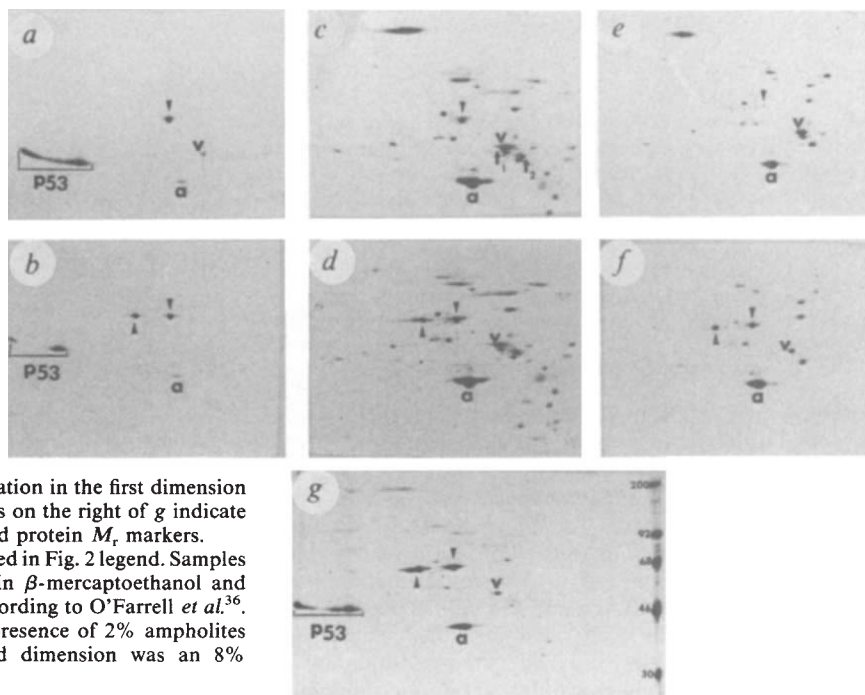
associated with p53 throughout the gradient, and the ratio between the two polypeptides appeared to be constant. Unlike p53, a substantial fraction of HSP70 was in monomers or low-*M_r* oligomers (Fig. 4b).

In conclusion, p53 can associate specifically with at least two rat HSPs, HSP68 and HSP70. A similar interaction seems to exist in transformed mouse cells that overproduce p53 because of a naturally occurring event rather than to *in vitro* transfection^{13,14}. It is therefore plausible that this complex is of physiological importance and may even be related to the transformation process.

One obvious concern is that the p53–HSP70 complex may be an extraction artefact. Although this is hard to disprove, the inability to detect a comparable complex in heat-shocked Rat-1 cells (data not shown) strongly militates against it. Interestingly, a specific association has also been shown between another heat shock protein, HSP90, and pp60^{src} (ref. 27). It is tempting to speculate that this is not mere coincidence, but rather may reflect a more general involvement of HSPs in the regulation of cell proliferation.

Fig. 3 Two-dimensional polyacrylamide gel comparison of heat shock proteins and p53-co-precipitating polypeptides. Clone 6 and Rat-1 cells were either maintained at 37 °C (*a*, *c*, *e*) or exposed to 46 °C for 10 min (*b*, *d*, *f*, *g*). Clone 6 lysates containing equal amounts of acid-insoluble radioactivity (2×10^6 c.p.m.) were immunoprecipitated with monoclonal antibody PAb421 and analysed by two-dimensional gel electrophoresis (*a*, *b*). Aliquots of the radiolabelled total cell extracts (3.7×10^5 c.p.m.) were also taken directly for a similar analysis. *c*–Clone 6, 37 °C; *d*, clone 6, 46 °C; *e*, Rat-1, 37 °C; *f*, Rat-1, 46 °C; *g*, a mixture of an immunoprecipitate identical to that in *b* with an aliquot of unprocessed clone 6, 46 °C lysate (similar to that shown in *d* but representing only 35,000 acid-insoluble c.p.m.). Upward-pointing arrow, HSP68; downward-pointing arrow, HSP70; *a*, actin; *t*₁, α -tubulin; *t*₂, β -tubulin; *v*, vimentin. Direction of migration in the first dimension (isoelectric focusing) was from left to right. Numbers on the right of *g* indicate the positions and sizes ($\times 10^{-3}$) of co-electrophoresed protein *M_r* markers.

Methods. Cells were labelled and processed as described in Fig. 2 legend. Samples to be analysed were made 9.5 M in urea and 2% in β -mercaptoethanol and subjected to two-dimensional gel electrophoresis according to O'Farrell *et al.*³⁶. The first dimension used isoelectrofocusing in the presence of 2% ampholites (1.6% pH 5.0–7.0, 0.4% pH 3.5–10.0). The second dimension was an 8% SDS-polyacrylamide gel.



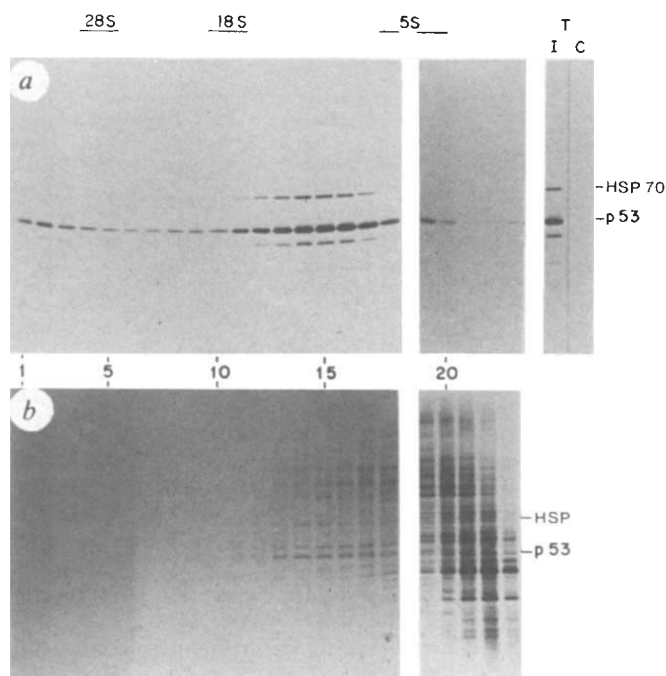


Fig. 4 Co-sedimentation of p53 and HSP70 on a sucrose gradient. A labelled cellular extract was fractionated on a sucrose gradient and each fraction was assayed for the presence of p53 and co-precipitating HSP70. *a*, Polypeptides immunoprecipitated with an anti-p53 monoclonal antibody. Fraction numbers are shown at the bottom; the positions of ribosomal RNA markers, sedimented in a parallel tube, are shown on top. T represents the proteins precipitated from an aliquot of the total (unfractionated) extract by either control (C) or anti-p53 monoclonal antibodies (I). *b*, Electrophoretic pattern of aliquots of each fraction subjected directly to gel analysis. The identification of HSP70 and p53 is tentative. Note that fractions 19–23 were electrophoresed on a separate gel and that the migration distances of individual polypeptides differed slightly between the two gels; we attempted to align the gels at the tentative position of the p53 band. The bands migrating ahead of p53 probably represent proteolytic cleavage products of this protein, present in variable quantities in different experiments and generated at least in part during the immunoprecipitation process³⁷. **Methods.** A 60-mm dish of clone 6 cells was labelled at 37 °C with 200 μ Ci ³⁵S-methionine for 4 h, extracted (see Fig. 1 legend) and sedimented through a 4.7-ml 5–20% (w/v) sucrose gradient containing 0.14 M NaCl, 10 mM Tris-Cl pH 8, 10 mM dithiothreitol, 1% aprotinin and 300 μ g ml⁻¹ phenylmethylsulphonyl fluoride, including a 0.4-ml 60% sucrose pad. Sedimentation was in a Beckman SW 50.1 rotor at 48,000 r.p.m. for 3 h at 4 °C. Fractions (0.21 ml, 23 fractions per gradient) were collected from the bottom of the tube following puncturing with a needle. Fraction 1 represents the fastest-sedimenting material. Aliquots (10 μ l) were mixed with sample buffer and directly analysed on a 12.5% SDS-polyacrylamide gel (*b*). The rest of the material underwent immunoprecipitation with monoclonal antibody PAb421 and was analysed similarly (*b*). Rat rRNA was sedimented through a parallel gradient and the positions of the respective size species determined using a spectrophotometer and confirmed by electrophoresis on a formaldehyde-agarose gel. Autoradiographic exposure time for both *a* and *b* was 1.5 h.

We have been unable to detect similar complexes convincingly in normal cells. In heat-induced cells both these HSPs are predominantly nuclear^{28,29}. If a nuclear site is necessary for the specific association, the apparent absence of p53 from nuclei of normal cells³⁰ may account for the inability of a complex to form. The nuclear accumulation of p53 in a variety of transformed cells³⁰ may thus lead to complex formation. Alternatively, the association between p53 and HSP70 may be a normal process, except that in such conditions the complex is very unstable and accumulates only in abnormal situations. Such abnormal conditions might simply be the overproduction of

p53, or structural changes in either HSP70 or p53 resulting in a more stable complex. The latter may pertain to Meth A cells, which contain strongly elevated p53 levels while possessing only twice the amount of the corresponding messenger RNA¹². Such cells exhibit a p53–HSP70 complex^{13,15}, and a relatively stable p53^{12,13}, and indeed seem to carry a mutated p53 gene encoding an altered protein product³⁸. It is not known whether this mutant p53 has a higher affinity for HSP70.

Stabilization of p53 in transformed cells may involve complex formation with HSP70, although it is equally conceivable that accumulation of the complex may in fact be a consequence of the stabilization. In this respect, the SV40 large-T antigen could be mimicking the action of a normal cellular protein, HSP70. This is also suggested by studies using the expression of mouse p53 in SV40-transformed monkey COS cells¹⁵, where anti-p53 monoclonal antibodies brought down both complexed large-T antigen and a ~70,000-*M_r* doublet, presumably HSP68 and HSP70. However, antibodies to SV40 large-T antigen efficiently co-precipitated only p53, with very little evidence for HSPs, suggesting that HSP70 and large-T antigen may be competing for the same or very close sites on the p53 molecule. Other similarities exist between HSP70 and the SV40 large-T antigen. Aside from having a nuclear affinity^{28,29,31}, both proteins are able to bind ATP^{24,31}. The SV40 large-T antigen possesses ATPase activity³¹, and although no activity has been demonstrated for HSP70²⁴, it is exhibited by the structurally related *Escherichia coli* DNA K protein^{32,33}. Interestingly, this bacterial protein is involved in the replication of bacteriophage DNA³³, as is large-T antigen in the replication of SV40³¹. Finally, clone 6 possesses a markedly elevated level of HSP70 without heat induction. The relationship between this apparent overproduction of HSP70 and the specific oncogenes activated in clone 6 remains unknown.

The data presented here raise the possibility that HSP70 is involved in cell proliferation, which is in keeping with the apparent cell-cycle-dependent expression of the human *hsp70* gene(s)³⁴. However, the significance of these findings can only be assessed when more is known about the molecular functions of HSP70.

We thank M. Tainovich for technical assistance and Drs O. Bensaude and B. Geiger for stimulating discussions. This work was supported by USPHS grant R01 CA 40099-01 from the National Cancer Institute, DHHS, and by a grant from the USA–Israel Binational Science Foundation. M.O. is a Scholar of the Leukemia Society of America, Inc.

Received 14 November 1985; accepted 13 January 1986.

1. Eliyahu, D., Raz, A., Gruss, P., Givol, D. & Oren, M. *Nature* **312**, 646–649 (1984).
2. Parada, L. F., Land, H., Weinberg, R. A., Wolf, D. & Rotter, V. *Nature* **312**, 649–651 (1984).
3. Jenkins, J. R., Rudge, K. & Currie, G. A. *Nature* **312**, 651–654 (1984).
4. Crawford, L. V. *et al.* *Cold Spring Harb. Symp. quant. Biol.* **44**, 179–187 (1980).
5. McCormick, F. & Harlow, E. *J. Virol.* **34**, 213–224 (1980).
6. Lane, D. P., Gannon, J. & Winchester, K. *Adv. viral Oncol.* **2**, 23–39 (1982).
7. Oren, M., Maltzman, W. & Levine, A. J. *Molec. cell. Biol.* **1**, 101–110 (1981).
8. Mora, P. T., Chandrasekaran, K., Hoffman, J. C. & McFarland, V. W. *Molec. cell. Biol.* **2**, 763–771 (1982).
9. May, E., Lasne, C., Prives, C., Borde, J. & May, P. J. *J. Virol.* **45**, 901–913 (1983).
10. Montenarh, M., Kohler, M., Aggeler, G. & Henning, R. *EMBO J.* (in the press).
11. Oren, M., Reich, N. C. & Levine, A. J. *Molec. cell. Biol.* **2**, 443–449 (1982).
12. Reich, N. C., Oren, M. & Levine, A. J. *Molec. cell. Biol.* **3**, 2143–2150 (1983).
13. Gronostajski, R. M., Goldberg, A. L. & Pardee, A. B. *Molec. cell. Biol.* **4**, 442–448 (1984).
14. Ruscetti, S. K. & Scolnick, E. M. *J. Virol.* **46**, 1022–1026 (1983).
15. Pinhasi, O. & Oren, M. *Molec. cell. Biol.* **4**, 2180–2186 (1984).
16. Harlow, E., Crawford, L. V., Pim, D. C. & Williamson, N. M. *J. Virol.* **39**, 861–869 (1981).
17. Rotter, V., Witte, O. N., Coffman, R. & Baltimore, D. *J. Virol.* **36**, 547–555 (1980).
18. Crawford, L. V. *Adv. viral Oncol.* **2**, 3–21 (1982).
19. Rotter, V., Friedman, H., Katz, A., Zerivitz, K. & Wolf, D. *J. Immun.* **131**, 329–333 (1983).
20. Cleveland, D., Fisher, S., Kirchner, M. & Laemmli, U. K. *J. biol. Chem.* **252**, 1102–1106 (1977).
21. Bensaude, O. & Morange, M. *EMBO J.* **2**, 173–177 (1983).
22. Morange, M., Diu, A., Bensaude, O. & Babinet, C. *Molec. cell. Biol.* **4**, 730–735 (1984).
23. Lowe, D. G. & Moran, L. A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2317–2321 (1984).
24. Welch, W. J. & Feramisco, J. R. *Molec. cell. Biol.* **5**, 1929–1937 (1985).
25. Steinberg, B., Pollack, R., Topp, W. & Botchan, M. *Cell* **13**, 19–32 (1978).
26. Eliyahu, D., Michalovitz, D. & Oren, M. *Nature* **316**, 158–160 (1985).
27. Opperman, H., Levinson, W. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1067–1071 (1981).
28. Welch, W. J. & Feramisco, J. R. *J. biol. Chem.* **259**, 4501–4513 (1984).
29. Velasquez, J. M. & Lindquist, S. *Cell* **36**, 655–662 (1984).
30. Rotter, V., Abutbul, H. & Ben-Zeev, A. *EMBO J.* **2**, 1041–1047 (1983).

31. Tooze, J. (ed.) *DNA Tumor Viruses* Pt 2 (Cold Spring Harbor Laboratory, New York, 1982).
32. Bardwell, J. C. A. & Craig, E. *Proc. natn. Acad. Sci. U.S.A.* **81**, 848-852 (1984).
33. Zylitz, M. & Georgopoulos, C. P. *J. biol. Chem.* **259**, 8820-8825 (1984).
34. Kao, H. T., Capassa, D., Heintz, N. & Nevins, J. R. *Molec. cell. Biol.* **5**, 628-633 (1985).
35. Curran, T., Van Beveren, C., Ling, N. & Verma, I. M. *Molec. cell. Biol.* **5**, 167-172 (1985).
36. O'Farrell, P. Z., Goodman, H. M. & O'Farrell, P. H. *Cell* **12**, 1133-1142 (1977).
37. Oren, M. & Levine, A. J. *Proc. natn. Acad. Sci. U.S.A.* **80**, 56-59 (1983).
38. Bienz, B., Zakut-Houri, R., Givol, D. & Oren, M. *EMBO J.* **3**, 2179-2184 (1984).
39. Mattzman, W., Oren, M. & Levine, A. J. *Virology* **112**, 145-156 (1981).

Product of *per* locus of *Drosophila* shares homology with proteoglycans

F. Rob Jackson*, Thaddeus A. Bargiello,
Suk-Hyeon Yun & Michael W. Young†

The Rockefeller University, 1230 York Avenue, New York,
New York 10021, USA

Genes controlling biological rhythms have been identified in *Drosophila*^{1,2}. The best characterized of these genes is called *period* (*per*). Although wild-type flies have daily (circadian) rhythms with a periodicity of ~24 h, *per*^s and *per*^l mutants have 19-h and 29-h rhythms, respectively, and *per*⁰ mutants are arrhythmic¹. The *per*^s mutation also enhances the sensitivity of the circadian clock to resetting by light stimuli³, and all three types of *per* mutations affect a much shorter period ultradian rhythm, the 55-s rhythm of the *Drosophila* courtship song⁴. A fragment of DNA of ~7 kilobases (kb) encoding a 4.5-kb poly(A)⁺ RNA restores rhythmicity when transduced into *Drosophila* carrying mutations^{5,6} or chromosomal deletions⁷ of the *per* locus. Here we report the sequence of this biologically active segment of DNA. The transcription unit that encodes the 4.5-kb RNA has been mapped, permitting a conceptual translation of a protein of 1,127 amino acids. Several abnormal phenotypes characterized by long-period rhythms are associated with changes in the sequence of untranslated portions of the transcription unit. The structure of some segments of the predicted protein suggests that it is a proteoglycan.

The structure of the 4.5-kb transcript was analysed by ribonuclease mapping of a series of RNA/RNA heteroduplexes. For each experiment, RNA complementary to a predetermined region of the 4.5-kb transcript was synthesized *in vitro* from a corresponding DNA segment of known sequence, cloned in pSP64 or pSP65 (ref. 7). Heteroduplexes were digested with RNase T₂ or a combination of RNase T₁ and RNase A, and resistant material was denatured and size-separated on acrylamide sequencing gels. Figure 1 shows the structures of most of the synthetic RNAs used to map the 4.5-kb transcript. Figure 2a-d shows the results of several of the mapping experiments.

To confirm the proposed structure of the 5' RNA coding region, two primer extension experiments were performed. The sequences of the two chemically synthesized primers, P1 and P2, are shown in Fig. 1. P1 corresponds to a DNA sequence that begins in RNA coding region 1, 81 base pairs (bp) from the putative 5' terminus, and P2 includes a sequence that starts in coding region 1, 167 bp from this expected terminus. Both primer extension experiments map the 5' end of the transcription unit to position 179 of Fig. 1. The result of the P1 primer extension experiment is shown in Fig. 2e.

Complementary DNA clones were produced by oligo(dT) priming of total adult poly(A)⁺ RNA as described in Fig. 3 legend. A 1.7-kb cDNA chosen for sequence analysis includes a poly(A) segment at its 3' terminus, allowing the 3' end of the transcription unit and a polyadenylation signal (position 7,370) to be identified (see Fig. 3, and legend). The cDNA contains most of coding region 5, beginning at position 5,524 of Fig. 3, and all of coding regions 6-8, as indicated at the top of Fig. 1.

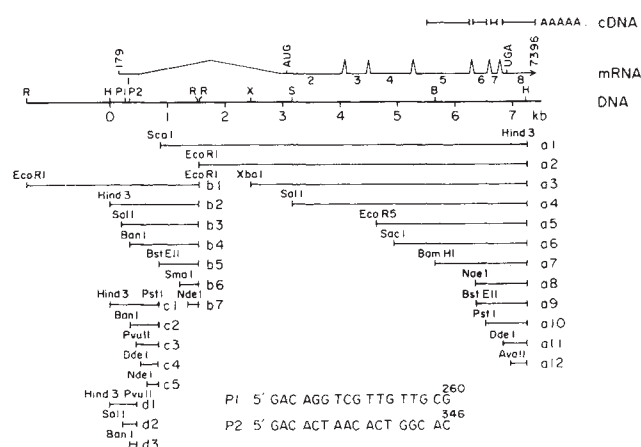


Fig. 1 Transcription map of the *per* locus. The structures of the restriction fragments cloned in pSP64 or pSP65 Riboprobe vectors (Promega Biotec) are indicated in the bottom half of the figure. For each fragment, ³²P-labelled, single-stranded RNA was synthesized from plasmid DNA after digestion with the restriction enzyme used to form the left end of each region depicted. Synthetic RNAs a1-12 extend leftward from a HindIII site, while b1-7, c1-5 and d1-3 extend leftward from EcoRI, PstI and PvuII sites, respectively. P1 and P2 represent the sequences and locations on the transcription unit of two primers used to map the transcription start site. The positions of transcription initiation (179) and termination (7,396) and of translation initiation (AUG) and termination (UGA) are indicated. For the restriction map in the centre of the figure, H = HindIII, R = EcoRI, X = XbaI, S = SalI and B = BamHI.

A comparison of this cDNA sequence with that of cloned genomic *per* DNA (presented below) has allowed us to locate several splice junctions, all of which agree with those established by ribonuclease mapping experiments. The cDNA sequence contains a single open reading frame that encodes 391 amino acids.

Figure 3 shows the DNA sequence of the interval containing the transcription unit that encodes the 4.5-kb *per* RNA. The length of the transcription unit is 7.217 kb. Also shown in Fig. 3 are the boundaries of the eight RNA coding regions which form the 4.5-kb messenger RNA. For those coding regions located only by ribonuclease mapping, splice junctions are predicted according to the consensus sequence rules of Mount⁸. The predicted size of the mRNA encoded by this transcription unit (minus a poly(A) tail) is 4.290 kb; this is in good agreement with a size of 4.5 kb determined previously by comparison with DNA and RNA size standards after electrophoresis of poly(A)⁺ RNA on formaldehyde/agarose gels^{9,10}.

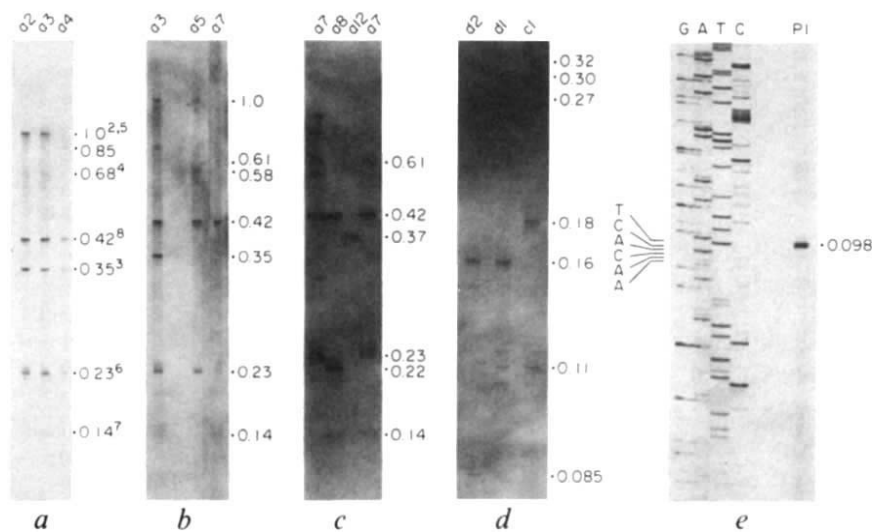
Figure 3 shows a conceptual translation of the spliced mRNA. Accordingly, RNA coding region 1 is not translatable. Translation is initiated near the beginning of RNA coding region 2, 2,909 bp from the beginning of the transcription unit (position 3,088; see also map in Fig. 1). All open reading frames are blocked upstream of this position. Translation could be initiated at the methionine codon at position 3,050 but a protein beginning at this position would be terminated 39 bp downstream (position 3,089). Translation of the predicted *per* protein terminates near the beginning of coding region 8 (position 6,891), so that 505 untranslated bases are found at the 3' end of the mRNA. The predicted *per* protein contains 1,127 amino acids.

A sequence analysis of the *per* locus allows a better understanding of a number of irregular clock phenotypes. A transcription study has been reported previously for several *per* locus mutations^{9,10}. Among these was a mutation associated with a chromosome translocation, *T(1;4)JC43*. Flies carrying this mutation have long-period rhythms (35 h), or can be arrhythmic^{11,12}. The 4.5-kb *per* transcript is absent in *T(1;4)JC43* flies, but a new transcript, of 11.5 kb, is synthesized in its place^{9,10}. Previous studies mapped the breakpoint of the translocation to

* Present address: Worcester Foundation for Experimental Biology, 222 Maple Avenue, Shrewsbury, Massachusetts 01545, USA.

† To whom reprint requests should be addressed.

Fig. 2 Mapping of *per* RNA coding regions. Labels at tops of panels *a-d* refer to transcribed segments depicted in Fig. 1. ³²P-labelled synthetic RNAs were transcribed from regions of cloned *per* DNA as described in Fig. 1 legend, then hybridized to 5–10 µg of total poly(A)⁺ RNA from Oregon R adults at 45 °C for 8–12 h (56 °C for a1–12), and heteroduplexes were digested with ribonucleases (RNase T₂, 50–100 U ml⁻¹ or RNase A and T₁, 100 µg ml⁻¹ and 10 U ml⁻¹) at 37 °C for 1 h. Protected fragments were denatured in formamide and separated on acrylamide gels as indicated in the text. Numbers indicate sizes (in kb) of ribonuclease-resistant RNA measured against single-stranded RNAs of known length transcribed from Riboprobe vectors. Superscripts to these numbers in *a* indicate the RNA coding region. Three fragments, of 0.32, 0.18 and 0.11 kb, are protected with synthetic RNAs b1 and c1 (*d*, lane 3 shows the results with c1); this seems to indicate that three RNA coding regions map to this interval. The two smaller fragments must be generated by secondary ribonuclease cleavage of the 0.32-kb fragment. RNA d1 (*d*, lane 2) protects all of the 0.11-kb region, but only part of the 0.32-kb region (0.30 kb) and part of the 0.18-kb region (0.16 kb). As both of the latter two regions are reduced to the same extent, they must end at the same position, about 20 bp to the right of the *Pvu*II site used to construct d1. This experiment also indicates that the 0.18-kb region is co-extensive with the right half of the 0.32-kb region. The 5' end of the 0.32-kb coding region must lie 0.30 kb to the left of the *Pvu*II site. RNA d2 protects most (0.085 kb) of the 0.11-kb region. The 5' end of the region, which is not protected, is therefore located ~25 bp to the left of a *Sall* site used to form RNA d2; this is 0.30 kb to the left of the *Pvu*II site used to map the boundaries of the 0.32-kb coding region above. As the 5' end of the 0.32-kb coding region is also 0.30 kb to the left of this *Pvu*II site, both regions begin at the same position and are co-extensive for 0.11 kb. The site of secondary cleavage proposed for the 0.32-kb coding region is composed of 18 consecutive A-T base pairs (see also Fig. 3). This sequence probably forms a hypersensitive substrate for the ribonucleases used in this mapping procedure. For *e*, ³²P-end-labelled primers (1–5 ng) were hybridized with total poly(A)⁺ RNA (10 µg) for 2–4 h at 42 °C and extended for 1 h with avian myeloblastosis virus (AMV) reverse transcriptase (50 U; BRL). DNA sequence ladders were produced by hybridizing the primers to cloned *per* DNA, followed by primer extension according to the dideoxy sequencing method^{24,25}.



a *Bgl*II/*Hind*III fragment⁹. A sequence analysis of the gene now allows the boundaries of the affected restriction fragment to be mapped to positions 7,217 and 7,234 respectively; this places the breakpoint of the translocation within the 3'-untranslated region of the transcription unit. Approximately 0.17 kb of the 4.5-kb mRNA has been exchanged, because of the translocation, for about 7 kb of RNA transcribed from chromosome 4 sequences. Ribonuclease mapping experiments have shown that sequences upstream of this breakpoint are spliced in the same way as wild-type 4.5-kb RNA (T.A.B., unpublished results). Presumably, *T(1;4)JC43* flies can still produce biological rhythms because a full-length *per* protein can be translated from the 11.5-kb fusion transcript.

Long-period rhythms are also generated by some transformed flies^{5,6}. In one study an attempt was made to construct transformants with a complete *per* transcription unit⁵. It is now apparent that these flies received DNA lacking 157 bp of the 3'-untranslated region. A 4.5-kb RNA is still produced by these long-period transformants, but at levels much lower than those found in wild-type flies⁵. The transcripts must terminate in vector sequences used to generate the transforming DNA.

In a second transformation study⁶, long-period rhythms were produced in flies transformed with DNA lacking sequences at the 5' end of the *per* transcription unit. From the sequence data presented here, these flies must lack the 5'-untranslated RNA coding region, but still retain all of the protein-coding sequence. Thus, if transcripts are produced by this DNA, a wild-type *per* protein would still be the expected translation product.

Thus, several long-period phenotypes can be associated with changes in the *per* transcription unit. In each case the change in DNA sequence fails to affect the structure of the *per* protein. Therefore, these long-period rhythms are probably generated by altered regulation of *per* protein synthesis.

Nearly half of the predicted *per* protein (47%) is composed of the five amino acids serine, threonine, glycine, alanine and proline, and these often form simple sequence tracts. For example, polyalanine and polyglycine tracts up to 17 residues long are found at positions 160, 801, 909 and 938. Given this

predominance of a small group of uncharged polar and nonpolar amino acids, the procedures of Kyte and Doolittle¹³ and Eisenberg *et al.*¹⁴ were used to perform a hydropathicity analysis of the protein. The protein contains no clearly hydrophobic regions, rather most protein segments have hydrophobicities and hydrophobic moment values that are characteristic of soluble proteins¹⁴.

The predicted *per* protein includes a potential site for phosphorylation by cyclic AMP-dependent protein kinase. This amino-acid sequence (Lys-His-Arg-Lys-Leu-Lys-Ser-Met) occurs near the carboxy terminus (residues 1,099–1,106) and resembles the phosphorylatable site in troponin I from rat skeletal muscle¹⁵. A synthetic peptide containing this sequence has been produced, and it appears to be an efficient substrate for phosphorylation *in vitro* (G. P. Gasic, unpublished). Various chemical agents commonly used to alter cyclic nucleotide levels induce a phase shift in circadian clocks^{16,17}. Thus, agents that affect cyclic nucleotide levels could alter the phosphorylation of a *per* protein.

The predicted *per* protein sequence has been compared with all sequences currently available in the Dayhoff and Doolittle protein sequence libraries. Using the FASTP computer program and the algorithm of Lipman and Pearson¹⁸, we detected extensive regions of homology to the core protein of a rat chondroitin sulphate proteoglycan. The bases for these homologies are the alternating Ser-Gly and Thr-Gly repeats of the *per* protein (residues 592–670). DNAs homologous to these unusual protein-coding segments have been found in vertebrates¹⁹. The protein core of the homologous chondroitin sulphate proteoglycan, which is 104 amino acids long, contains 49 strictly alternating Ser-Gly residues. This repetitive amino-acid sequence forms a substrate for the extensive glycosylation of the core protein²⁰. A partial analysis of the glycosylated region of a second proteoglycan, a rat heparin proteoglycan, has shown that it, too, is composed of only serine and glycine residues, many or all of which must form an alternating repeat²¹. Synthetic peptides containing alternating Ser-Gly residues behave as substrates for glycosyl transferases *in vitro*²².

71
 182
 213
 284
 305
 426
 497
 3053
 3114
 3168
 3222
 3276
 3330
 3384
 3438
 3492
 3546
 3600
 3654
 3708
 3762
 3816
 3870
 3924
 3978
 4032
 4099
 4153
 4207
 4261
 4315
 4369
 4423
 4482
 4542
 4596
 4650
 4704
 4758
 4812
 4866
 4920

570
 588
 606
 624
 642
 660
 678
 696
 714
 732
 750
 768
 786
 804
 822
 840
 858
 876
 894
 912
 930
 948
 966
 984
 1002
 1020
 1038
 1056
 1074
 1092
 1110
 1128
 1146
 1164
 1182
 1200
 1218
 1236
 1254
 1272
 1290
 1308
 1326
 1344
 1362
 1380
 1398
 1416
 1434
 1452
 1470
 1488
 1506
 1524
 1542
 1560
 1578
 1596
 1614
 1632
 1650
 1668
 1686
 1704
 1722
 1740
 1758
 1776
 1794
 1812
 1830
 1848
 1866
 1884
 1902
 1920
 1938
 1956
 1974
 1992
 2010
 2028
 2046
 2064
 2082
 2100
 2118
 2136
 2154
 2172
 2190
 2208
 2226
 2244
 2262
 2280
 2298
 2316
 2334
 2352
 2370
 2388
 2406
 2424
 2442
 2460
 2478
 2496
 2514
 2532
 2550
 2568
 2586
 2604
 2622
 2640
 2658
 2676
 2694
 2712
 2730
 2748
 2766
 2784
 2802
 2820
 2838
 2856
 2874
 2892
 2910
 2928
 2946
 2964
 2982
 3000
 3018
 3036
 3054
 3072
 3090
 3108
 3126
 3144
 3162
 3180
 3198
 3216
 3234
 3252
 3270
 3288
 3306
 3324
 3342
 3360
 3378
 3396
 3414
 3432
 3450
 3468
 3486
 3504
 3522
 3540
 3558
 3576
 3594
 3612
 3630
 3648
 3666
 3684
 3702
 3720
 3738
 3756
 3774
 3792
 3810
 3828
 3846
 3864
 3882
 3900
 3918
 3936
 3954
 3972
 3990
 4008
 4026
 4044
 4062
 4080
 4098
 4116
 4134
 4152
 4170
 4188
 4206
 4224
 4242
 4260
 4278
 4296
 4314
 4332
 4350
 4368
 4386
 4404
 4422
 4440
 4458
 4476
 4494
 4512
 4530
 4548
 4566
 4584
 4602
 4620
 4638
 4656
 4674
 4692
 4710
 4728
 4746
 4764
 4782
 4800
 4818
 4836
 4854
 4872
 4890
 4908
 4926
 4944
 4962
 4980
 4998
 5016
 5034
 5052
 5070
 5088
 5106
 5124
 5142
 5160
 5178
 5196
 5214
 5232
 5250
 5268
 5286
 5304
 5322
 5340
 5358
 5376
 5394
 5412
 5430
 5448
 5466
 5484
 5502
 5520
 5538
 5556
 5574
 5592
 5610
 5628
 5646
 5664
 5682
 5700
 5718
 5736
 5754
 5772
 5790
 5808
 5826
 5844
 5862
 5880
 5898
 5916
 5934
 5952
 5970
 5988
 6006
 6024
 6042
 6060
 6078
 6096
 6114
 6132
 6150
 6168
 6186
 6204
 6222
 6240
 6258
 6276
 6294
 6312
 6330
 6348
 6366
 6384
 6402
 6420
 6438
 6456
 6474
 6492
 6510
 6528
 6546
 6564
 6582
 6600
 6618
 6636
 6654
 6672
 6690
 6708
 6726
 6744
 6762
 6780
 6798
 6816
 6834
 6852
 6870
 6888
 6906
 6924
 6942
 6960
 6978
 6996
 7014
 7032
 7050
 7068
 7086
 7104
 7122
 7140
 7158
 7176
 7194
 7212
 7230
 7248
 7266
 7284
 7302
 7320
 7338
 7356
 7374
 7392
 7410
 7428
 7446
 7464
 7482
 7500
 7518
 7536
 7554
 7572
 7590
 7608
 7626
 7644
 7662
 7680
 7698
 7716
 7734
 7752
 7770
 7788
 7806
 7824
 7842
 7860
 7878
 7896
 7914
 7932
 7950
 7968
 7986

Fig. 3 Sequence of genomic *per* DNA. The DNA sequence begins at coordinate 0 of the Fig. 1 map (a *Hind*III recognition sequence), and ends 91 bp downstream of the site of poly(A) addition, at coordinate 7,396. mRNA coding regions are indicated by upper case letters; intervening sequences, and 5' and 3' non-transcribed DNA, are indicated by lower case letters. A conceptual translation of the *per* protein is given above the DNA sequence, with gaps in the amino-acid sequence corresponding to the locations of intervening sequences. (The complete sequence of intervening sequence 1 is available from the authors on request.) **Methods.** Genomic DNA fragments were sequenced using the dideoxy method^{24,25}. Fragments suitable for shotgun sequencing were generated by digesting Canton S, *per* locus DNA with one of four restriction enzymes, each of which cleaves a 4-bp recognition sequence (*Sau*3A, *Alu*I, *Fnu*dII and *Rsa*I). These smaller fragments were randomly subcloned in M13 (mp18 or mp19), and the database of sequences obtained was assembled into longer contiguous DNA sequences using the programs of Staden²⁶. Gaps in the complete sequence of the genomic DNA were resolved by cloning selected subregions and generating ordered sets of overlapping deletions according to the method of Dale *et al.*²⁷. 27 kb of DNA was sequenced to generate the complete sequence for both strands of *per*. Thus, each base was read twice on average. cDNA libraries were constructed from 1–2 µg of poly(A)⁺ RNA isolated from adult Canton S *Drosophila melanogaster*. RNA was annealed to oligo(dT) primers which were elongated with AMV reverse transcriptase. Second-strand synthesis was done using the Klenow fragment of *Escherichia coli* DNA polymerase. The product was digested with *S*₁ nuclease, ends were repaired with DNA polymerase, and DNA was modified with *Eco*RI methylase. Synthetic *Eco*RI linkers were ligated, cut with *Eco*RI, and DNA fragments were size-selected on Bio-Gel A15 (Bio-Rad). DNA was cloned in λgt10. The *Eco*RI fragment from one *per*-homologous clone was ligated to M13mp18. Processive deletions were produced²⁷, and the complete DNA sequence was obtained by conventional methods^{24,25}.

Whereas serines are glycosylated in the Ser-Gly repeats of the two naturally occurring proteoglycans described above, threonine is glycosylated in other proteoglycans that are only just beginning to be characterized²³. The modification of proteoglycans containing Ser-Gly repeats occurs through an O-glycosylation of serine. Aside from serine, only threonine can be modified in this way²³. Thus, both of the unusual Ser-Gly and Thr-Gly repeat tracts in the *per* protein might serve as sites for glycosylation.

Although relatively little is known about the function of proteoglycans, their distribution has been well studied; they are found in extracellular locations or in association with cell surfaces. Given the similarities between *per* protein and known proteoglycans, this is where we would expect to find this clock protein and perhaps other components of the biological clock that it controls.

We thank Laurel Eckhardt, Simon Kidd, Peter Model and Norton D. Zinder for comments on the manuscript, and Peter Model for providing the synthetic oligonucleotide primers. This work was supported by grants to M.W.Y. from the NIH and the Andre and Bella Meyer Foundation.

Received 7 November; accepted 30 December 1985.

1. Konopka, R. J. & Benzer, S. *Proc. natn. Acad. Sci. U.S.A.* **68**, 2112-2116 (1971).

2. Jackson, F. R. *J. Neurogenet.* **1**, 3-15 (1983).
3. Konopka, R. J. *Fedn Proc.* **38**, 2602-2605 (1979).
4. Kyriacou, C. P. & Hall, J. C. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6729-6733 (1980).
5. Bargiello, T. A., Jackson, F. R. & Young, M. W. *Nature* **312**, 752-754 (1984).
6. Zehring, W. A. *et al. Cell* **39**, 369-376 (1984).
7. Melton, D. A. *et al. Nucleic Acids Res.* **12**, 7035-7056 (1984).
8. Mount, S. M. *Nucleic Acids Res.* **10**, 459-472 (1982).
9. Bargiello, T. A. & Young, M. W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2142-2146 (1984).
10. Reddy, P. *et al. Cell* **38**, 701-710 (1984).
11. Young, M. W. & Judd, B. H. *Genetics* **88**, 723-742 (1978).
12. Smith, R. F. & Konopka, R. J. *Molec. gen. Genet.* **183**, 243-251 (1981).
13. Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105-132 (1982).
14. Eisenberg, D., Schwarz, E., Komaromy, M. & Wall, R. *J. molec. Biol.* **179**, 125-142 (1984).
15. Bramson, H. N., Kaiser, E. T. & Mildvan, A. S. *CRC Crit. Rev. Biochem.* **15** (No. 2), 93-124 (1984).
16. Eskin, A., Corrent, G., Lin, C.-Y. & McAdoo, D. J. *Proc. natn. Acad. Sci. U.S.A.* **79**, 660-664 (1982).
17. Eskin, A. & Takahashi, J. S. *Science* **220**, 82-84 (1983).
18. Lipman, D. J. & Pearson, W. R. *Science* **227**, 1435-1441 (1985).
19. Shin, H. S., Bargiello, T. A., Clark, B. T., Jackson, F. R. & Young, M. W. *Nature* **317**, 445-448 (1985).
20. Bourdon, M. A., Oldberg, A., Pierschbacher, M. & Ruoslahti, E. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1321-1325 (1985).
21. Robinson, H. M., Horner, K. A., Hook, M., Ogren, S. & Lindahl, U. *J. biol. Chem.* **253**, 6687-6693 (1978).
22. Coudron, C., Loerner, T., Jacobson, I., Roden, L. & Schwartz, N. B. *Fedn Proc. Abstr.* **39**, 1671 (1980).
23. Kornfeld, R. & Kornfeld, S. in *The Biochemistry of Glycoproteins and Proteoglycans* (ed. Lennarz, W. J.) 1-34 (Plenum, New York, 1980).
24. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
25. Biggin, M. D., Gibson, T. J. & Hong, G. F. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3963-3965 (1983).
26. Staden, R. *Nucleic Acids Res.* **10**, 4731-4751 (1982).
27. Dale, R. M. K., McClure, B. A. & Houchins, J. P. *Plasmid* **13**, 31-40 (1985).

Existence of distinct sodium channel messenger RNAs in rat brain

Masaharu Noda, Takayuki Ikeda, Toshiaki Kayano, Harukazu Suzuki, Hiroshi Takeshima, Mika Kurasaki, Hideo Takahashi & Shosaku Numa

Departments of Medical Chemistry and Molecular Genetics, Kyoto University Faculty of Medicine, Kyoto 606, Japan

The sodium channel is a voltage-gated ionic channel essential for the generation of action potentials¹⁻³. It has been reported that the sodium channels purified from the electric organ of *Electrophorus electricus* (electric eel)^{4,5} and from chick cardiac muscle⁶ consist of a single polypeptide of relative molecular mass (M_r) ~260,000 (260K), whereas those purified from rat brain⁷ and skeletal muscle⁸ contain, in addition to the large polypeptide, two or three smaller polypeptides of M_r 37-45K. Recently, we have elucidated the primary structure of the *Electrophorus* sodium channel by cloning and sequencing the DNA complementary to its messenger RNA⁹. Despite the apparent homogeneity of the purified sodium channel preparations, several types of tetrodotoxin (or saxitoxin) binding sites or sodium currents have been observed in many excitable membranes¹⁰⁻¹⁹. The occurrence of distinguishable populations of sodium channels may be attributable to different states of the same channel protein or to distinct channel proteins. We have now isolated complementary DNA clones derived from two distinct rat brain mRNAs encoding sodium channel large polypeptides and present here the complete amino-acid sequences of the two polypeptides (designated sodium channels I and II), as deduced from the cDNA sequences. A partial DNA sequence complementary to a third homologous mRNA from rat brain has also been cloned.

Figure 1 shows the nucleotide sequences of the cloned cDNAs encoding rat brain sodium channels I and II (for the cloning procedure, see Fig. 1 legend). The primary structures of the two sodium channel large polypeptides were deduced by using the reading frame corresponding to the amino-acid sequence of the *Electrophorus* counterpart. The assignment of the translation initiation sites was based on the fact that they represent the first ATG triplets that appear downstream of a nonsense codon found in-frame. Rat brain sodium channels I and II are composed of

2,009 (or 1,998 with the deletion in clone prSCH28) and 2,005 amino acids, with calculated M_r s of 228,758 (227,576) and 227,840, respectively; at positions where amino-acid differences are predicted by the nucleotide differences found among the individual clones used (see Fig. 1 legend), the amino acids given in Fig. 1 have been used to calculate the M_r . Nucleotide sequence analysis of the third type of cloned cDNA (clone prSCH203; see Fig. 1 legend) revealed that it encodes a sequence of 573 amino acids that is highly homologous with the carboxy-terminal sequences of sodium channels I and II (data not shown).

The amino-acid sequences of rat sodium channels I and II are compared with that of *Electrophorus* sodium channel in Fig. 1. The degree of sequence homology is 87%, 62% and 62% for the rat I/rat II, rat I/*Electrophorus* and rat II/*Electrophorus* pairs, respectively. Rat sodium channel I as well as II, like the *Electrophorus* sodium channel⁹, contains four homologous internal repeats with deletions (or insertions) (positions 111-456, 775-1,047, 1,239-1,553 and 1,562-1,860 in the aligned sequences). The regions corresponding to these repeats are highly conserved among the three sodium channels, whereas the remaining regions, all of which are assigned to the cytoplasmic side of the membrane (see below), are less well conserved, except for the region between repeats III and IV. A large insertion of 166-178 amino acids occurs in the region between repeats I and II of the rat sodium channels, as compared with the *Electrophorus* counterpart. The inserted segment contains several potential sites of phosphorylation by cyclic AMP-dependent protein kinase²⁰⁻²², which are conserved in the two rat sodium channels.

Rat sodium channels I and II have hydropathicity profiles²³ similar to that of the *Electrophorus* sodium channel⁹ (data not shown). Each internal repeat has five hydrophobic segments (S1, S2, S3, S5 and S6) and one positively charged segment (S4), all of which exhibit predicted secondary structure²⁴. Segments S1, S2 and S3 generally contain a few charged residues. S3 has a few negative charges, and S2 a few negative and positive charges, whereas S1 is not uniform in charge. Segment S4 contains four to eight arginine or lysine residues located at every third position, with mostly nonpolar residues intervening between the basic residues. Segments S5 and S6 are highly hydrophobic regions without any charged residues (except S6 in repeat II of sodium channel I).

We have previously assumed that the four repeated units of homology in the sodium channel are oriented in a pseudosymmetric fashion across the membrane and that its amino terminus is located on the cytoplasmic side of the membrane⁹. This model

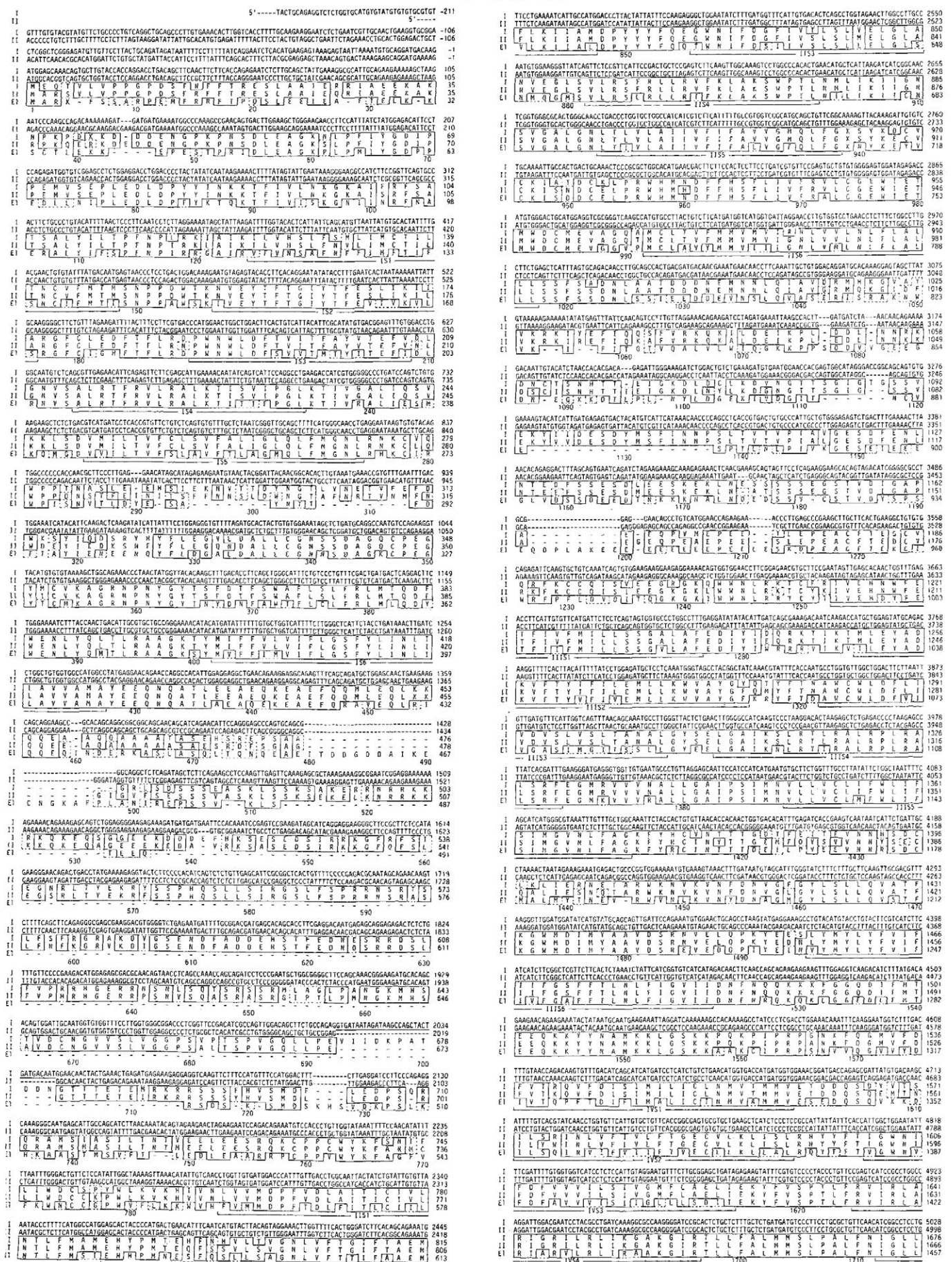


Fig. 1 Nucleotide sequences of cDNAs encoding rat sodium channels I (first row) and II (second row) and alignment of the deduced amino-acid sequences of rat sodium channels I (third row) and II (fourth row) and of the *E. electricus* sodium channel⁹ (fifth row) (see overleaf).

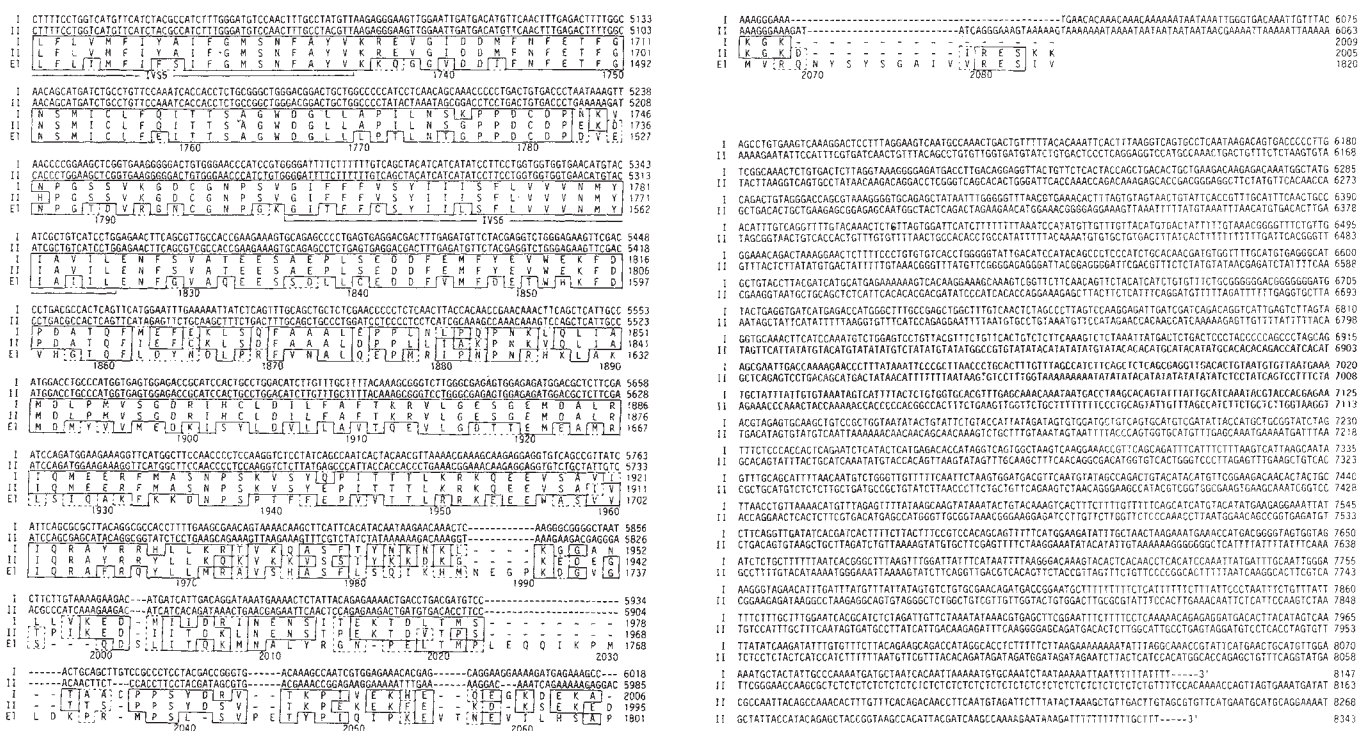


Fig. 1 Nucleotide sequences of cDNAs encoding rat sodium channels I (first row) and II (second row) and alignment of the deduced amino-acid sequences of rat sodium channels I (third row) and II (fourth row) and of the *E. electricus* sodium channel⁹ (fifth row). Nucleotide and amino-acid residues are numbered as in ref. 9; numbers of the last residues are given in the right-hand side. Identical amino-acid residues are boxed with solid lines and conservative amino-acid residues³⁵ with broken lines. Gaps (-) have been inserted into the amino-acid and nucleotide sequences to achieve maximum amino-acid sequence homology. Numbers of positions in the aligned amino-acid sequences into which gaps have been introduced are given beneath the sequences. For evaluating amino-acid sequence homology (see text), a continuous stretch of gaps has been counted as one substitution regardless of its length. Segments S1-S6 in each of repeats I-IV are indicated. The potential N-glycosylation sites^{30,31} conserved in all three sodium channels are the asparagine residues at positions 212, 285, 303, 340, 917, 1,417 and 1,431. Nucleotide 8,147 in prSCH109, nucleotide 7,682 in prSCH202 and nucleotide 8,343 in prSCH207 are followed by a poly(dA) tract which is connected with the vector DNA sequence³⁶. A polyadenylation signal, AATAAA (ref. 37) or ATTAAA (ref. 38), is found at nucleotides 8,125-8,130 for prSCH109 and at nucleotides 7,662-7,667 and 8,320-8,325 for prSCH202 and prSCH207, respectively. The nucleotide differences observed among the individual clones are as follows. Sodium channel I cDNA: insertion (prSCH71) or deletion (prSCH28) of nucleotides 2,011-2,043 (underlined); G (prSCH74) or A (prSCH71) at nucleotide 1,441; G (prSCH28) or A (prSCH71) at 1,780; G (prSCH28) or A (prSCH71) at 2,467. Sodium channel II cDNA: A (prSCH93) or G (prSCH9) at nucleotide 1,149; G (prSCH78) or A (prSCH9) at 1,511; C (prSCH78) or T (prSCH11) at 2,688; G (prSCH78 and prSCH82) or A (prSCH11) at 3,065; G (prSCH11 and prSCH82) or A (prSCH78) at 3,313; G (prSCH82) or A (prSCH11 and prSCH78) at 3,357; T (prSCH82) or C (prSCH202) at 3,669; C (prSCH82) or G (prSCH202) at 3,671; A (prSCH82) or G (prSCH202) at 3,684. The resulting amino-acid substitutions are as follows. Sodium channel I: insertion/deletion of residues 671-681; Asp or Asn at residue 481; Glu or Lys at 594; Asp or Asn at 823. Sodium channel II: Arg or Lys at residue 504; Arg or His at 1,022; Val or Met at 1,105; Ala or Gly at 1,224. Wherever a nucleotide or amino-acid difference occurs, the former residue is shown in the sequences presented. S₁ nuclease mapping³⁹ of rat brain poly(A)⁺ RNA with a probe containing the *Eco*RI (nucleotide 1,556)-*Bam*HI(2,116) segment of prSCH71 indicated the presence of the sodium channel I mRNA both with and without the deletion of nucleotides 2,011-2,043.

Methods. Poly(A)⁺ RNA was prepared^{40,41} from the whole brain of male Wistar rats (~200 g body weight). A cDNA library, constructed⁴² using 20 µg of oligo(dT)₁₂₋₁₈ and 20 µg of rat brain poly(A)⁺ RNA, was screened by hybridization at 50°C with an *E. electricus* sodium

channel cDNA probe, that is, an equimolar mixture of the *Hind*III(-250)-*Hind*III(1,777) fragment from prSCH411 and the *Hind*III(1,777)-*Ban*II(3,307) and *Ban*II(3,307)-*Xba*I(5,615) fragments from prSCH50 (ref. 9). Transformation and screening were done as described previously⁹. Of the 13 positive clones isolated from ~1 × 10⁵ transformants, four (prSCH7, prSCH8, prSCH9 and prSCH11) hybridized with an RNA species of ~9,000-9,500 nucleotides which their cDNA inserts were used as probes for blot hybridization analysis⁴³ of rat brain poly(A)⁺ RNA. The cDNA insert of prSCH7 was shown by restriction mapping to comprise part of that of prSCH8. Nucleotide sequence analysis⁴⁴ revealed that prSCH8, prSCH9 and prSCH11 carried cDNA sequences encoding amino-acid sequences homologous with different regions of the *E. electricus* sodium channel. Another cDNA library was constructed³⁰ using 12.6 µg of poly(A)⁺ RNA and 5.6 µg of the vector-primer DNA. From ~2 × 10⁵ transformants, five positive clones were isolated by hybridization at 60°C with an equimolar mixture of the *Hinf*I(4,935)-*Hinf*I(5,658) fragment from prSCH8 and the *Hinf*I(2,568)-*Hinf*I(3,332) fragment from prSCH11. Restriction mapping indicated that all five clones were derived from an identical mRNA species. Sequence analysis showed that the 5'-terminal 219-nucleotide sequence of prSCH109, which harboured the largest cDNA insert, was similar but not identical to the 3'-terminal sequence of the cDNA insert of prSCH11, indicating that the two clones were derived from distinct mRNA species. Screening of an additional ~6 × 10⁵ transformants from the same Okayama-Berg library by hybridization at 55°C with the *Pst*I(4,350)-*Pvu*II(5,498) fragment from prSCH109 yielded 14 positive clones. These clones were classified into three groups by restriction mapping: nine clones belonged to the same group as prSCH109, three (for example, prSCH202 and prSCH207) to another group, and two (for example, prSCH203) to the third group. For cloning cDNA sequences further upstream, the oligodeoxynucleotide primer 5'-TCTGTGTTTAAAGTTT-3', complementary both to nucleotides 3,375-3,389 in prSCH109 and to nucleotides 3,345-3,359 in prSCH11, was prepared using an automatic DNA-synthesizer (Applied Biosystems) and was elongated by reverse transcription. Single-stranded cDNA was synthesized using 1 nmol of this primer and 100 µg of rat brain poly(A)⁺ RNA and was converted to double-stranded cDNA, which was cloned in pBR322; the procedures were essentially the same as reported previously⁴², except that poly(A)⁺ RNA was denatured using methylmercuric hydroxide⁴⁵. From ~4 × 10⁴ transformants, 46 positive clones were isolated by hybridization at 55°C with an equimolar mixture of the *Hinf*I(1,144)-*Pst*I(1,396) fragment from prSCH9 and the *Bam*HI(2,440)-*Eco*RI(3,068) and *Eco*RI(3,068)-*Eco*RI(3,362) fragments from prSCH11. Restriction mapping of 23 of the clones indicated that 11 (for example, prSCH28, prSCH71 and prSCH74) were derived from the mRNA represented by prSCH109, and 9 clones (for example, prSCH78) from the mRNA represented by prSCH11. For isolating

postulates the presence of an even number of transmembrane segments in each repeat because no additional hydrophobic segments are predicted outside the repeats. Thus, we have proposed that either four or all six of segments S1–S6 in each repeat traverse the membrane⁹. If four transmembrane segments are assumed, it seems likely that segments S5 and S6 and two of segments S1, S2 and S3 span the membrane. This transmembrane topology would be favoured if each of the four positively charged segments S4 interacted with each of the four negatively charged segments located between segment S6 of repeat II and segment S1 of repeat III⁹. However, the clusters of acidic residues in this region are not as conspicuous in the rat sodium channels as in the *Electrophorus* counterpart. Furthermore, the voltage-dependent gating of the sodium channel implies the presence of a voltage sensor, which is thought to be a collection of charges or equivalent dipoles moving under the influence of the membrane electric field^{1,25}. In fact, this movement can be measured as a gating current²⁶. The finding that the equivalent of four to six charges must move fully across the membrane to open one sodium channel^{1,25} suggests the intramembranous location of many dipoles that move by smaller distances. These considerations favour the possibility that all of the six segments S1–S6 span the membrane (Fig. 2a), presumably forming α -helical structures^{27–29}. The ion selectivity of the sodium channel makes it unlikely that segment S4, which has many positive charges, constitutes the inner wall of the channel; it seems more reasonable for this segment to be localized within clusters of other segments (Fig. 2b). The formation of ion pairs between many of the positive charges in segment S4 and the negative charges in other segments such as S3 and S1 would stabilize the intramembranous location of the charged segments. The transmembrane topology shown in Fig. 2a is consistent with six of the seven potential *N*-glycosylation sites^{30,31} that are conserved in all three sodium channels (see Fig. 1 legend) as well as with all the potential cyclic AMP-dependent phosphorylation sites^{20–22}.

The unique structure of segment S4 in all repeats is strikingly well conserved among the three sodium channels. It seems probable that the positive charges present in this segment, many of which presumably form dipoles, represent the voltage sensor. They would move outward in response to depolarization, causing conformational changes and possible rearrangement of ion pairs. Segment S4 in repeats I, II, III and IV contains four, five, six and eight positive charges, respectively; this may underlie the proposal that the successive transitions among several closed states are fast and involve less charge movement than the slow transition from the last closed state to the open state³².

As the minimum cross-section of the sodium channel is thought to be $3 \times 5 \text{ \AA}$ (refs 33, 34), at least one helical segment

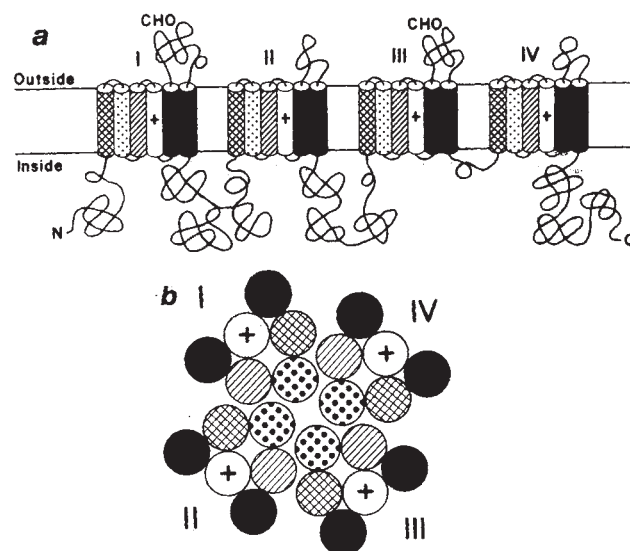


Fig. 2 a, Proposed transmembrane topology of the sodium channels; b, proposed arrangement of the transmembrane segments viewed in the direction perpendicular to the membrane. In a, the four units of homology spanning the membrane are displayed linearly. Segments S1–S6 in each repeat (I–IV) are indicated by cylinders as follows: S1, cross-hatched; S2, stippled; S3, hatched; S4, indicated by a plus sign; S5 and S6, solid. Putative sites of *N*-glycosylation (CHO) are indicated. In b, the ionic channel is represented as a central pore surrounded by the four units of homology. Segments S1–S6 in each repeat (I–IV) are represented by circles indicated as in a.

from each repeat would participate in forming the inner wall of the channel at the narrowest point. Remarkably, segment S2 in every repeat contains, at equivalent positions, a glutamic acid residue (positions 169, 839, 1,305 and 1,626 in the aligned sequences) and a lysine residue (positions 173, 843, 1,309 and 1,630), which are conserved in all three sodium channels. These acidic and basic residues are located on one side of the α -helix. In addition, segment S2 in repeats I and III contains a conserved glutamic acid or aspartic acid residue (positions 159 and 1,295) located 10 residues from the above-mentioned glutamic acid on nearly the same side of the α -helix; segment S2 in repeat III has two more conserved charged residues (glutamic acid at position 1,292 and lysine at position 1,296), which face another side of the α -helix. In view of the high degree of conservation of the charged residues in segment S2, it is tempting to hypothesize that this segment forms the inner wall of the channel (Fig. 2b).

Segment S3 in every repeat also contains a conserved aspartic acid residue (positions 195, 861, 1,327 and 1,647) at an equivalent position. It is an intriguing speculation that this aspartic acid, which may be paired with a basic residue in segment S4 in the closed state of the channel, may be shifted by depolarization to interact with the lysine present in segment S2 of every repeat; this would cause a conformational change of the putative channel wall and simultaneously mask the positive charge on the channel lining to allow sodium ions to pass through the central pore. Clusters of positively charged residues (predominantly lysine) and of negatively charged residues are conserved in the regions preceding segment S1 and following segment S6 of repeat IV, respectively. It is conceivable that these segments, which are assigned to the cytoplasmic side of the membrane, are involved in the inactivation of the sodium channel³².

Poly(A)⁺ RNA from adult rat brain, skeletal muscle and cardiac muscle was subjected to blot hybridization analysis with cDNA probes for rat brain sodium channels I and II (Fig. 3a, b). The brain contained RNA species of ~9,000 and ~9,500 nucleotides that were hybridizable with the sodium channel I

prSCH82 and prSCH93, the synthetic oligodeoxynucleotide primers 5'-TTCCCCCGTGGTGTA-3' (complementary to nucleotides 4,105–4,119 in prSCH202) and 5'-ACTGACGAAGCTCTCC-3' (complementary to nucleotides 1,452–1,466 in prSCH9 and prSCH78) were elongated using a mixture of 3 nmol of the former primer and 2 nmol of the latter primer and 200 μ g of rat brain poly(A)⁺ RNA (see above). Selection of the cloned cDNA transcripts ($\sim 4 \times 10^4$ transformants) by hybridization at 60 °C with the *Eco*RI(3,068)–*Eco*RI(3,362) fragment from prSCH11 and with the *Hinf*I(1,144)–*Pst*I(1,396) fragment from prSCH9 yielded two (for example, prSCH82) and five (for example, prSCH93) positive clones, respectively. DNA sequencing was done using the method of Maxam and Gilbert⁴⁴; both strands were sequenced, except for a small portion of the 3'-noncoding sequence (nucleotides 8,309–8,343) of the sodium channel II cDNA. The clones that were subjected to DNA sequencing were as follows. Sodium channel I: prSCH74 (carrying nucleotides –251 to 1,624), prSCH71 (1,245–2,709), prSCH28 (1,724–3,389 with a deletion of 2,011–2,043), prSCH109 (3,234–8,147), prSCH8 (4,561–6,039). Sodium channel II: prSCH93 (–210 to 1,458), prSCH9 (979–1,511), prSCH78 (1,336–3,359), prSCH11 (2,139–3,428), prSCH82 (2,983–4,119), prSCH202 (3,669–7,682), prSCH207 (4,234–8,343; nucleotides 4,290–7,553 were not sequenced). The restriction maps and sequencing strategy are available from the authors on request.

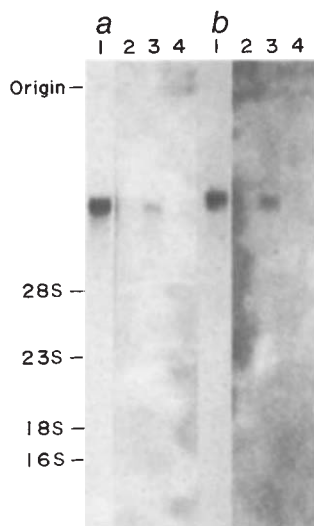


Fig. 3 Autoradiogram of blot hybridization analysis of poly(A)⁺ RNA from excitable tissues with sodium channel I (a) and II (b) cDNA probes. Poly(A)⁺ RNA (20 µg) from whole brain (lane 1), skeletal muscle (lane 2), cardiac muscle (lane 3) and liver (lane 4) of male Wistar rats (~200 g body weight) was denatured with 1 M glyoxal and 50% dimethyl sulphoxide, electrophoresed on a 1.2% agarose gel and transferred to a nitrocellulose filter⁴³. The filter was divided into two portions and subjected to hybridization with the *Hind*III(3,534)–*Hind*III(5,811) fragment from prSCH109 (a) or the *Sac*I(3,931)–*Eco*RI(5,869) fragment from prSCH202 (b). The probes were labelled by nick-translation⁴⁶ with [α -³²P]dCTP. Autoradiography was performed for 10 h (lane 1) or 72 h (lanes 2–4). The size markers used were rat and *Escherichia coli* ribosomal RNAs.

and II cDNA probes, respectively, in comparable amounts (Fig. 3, lanes 1). These sizes are consistent with those expected from the cloned cDNAs. The cardiac muscle also contained two RNA species that were hybridizable with the respective cDNA probes and indistinguishable in size from the brain mRNAs, although the hybridization signals were much weaker (Fig. 3, lanes 3); note the difference in the duration of autoradiography (see Fig. 3 legend). No hybridizable RNA species were detected with the skeletal muscle RNA (lane 2) or with the liver RNA used as a control (lane 4). The weak or undetectable hybridization observed with the cardiac muscle and skeletal muscle RNAs may be accounted for either by their low content of the sodium channel mRNAs or by the presence of different sodium channel mRNAs in these tissues. It has been reported that adult rat cardiac muscle contains a sodium channel species with a low affinity for tetrodotoxin, in addition to a species of sodium channel that, like the brain counterpart, has a high affinity for the toxin, and that the concentration of the high-affinity sodium channel is much lower than that of the low-affinity sodium channel^{10,11}. The structurally distinct sodium channels encoded by the mRNAs found may be responsible for the different types of sodium currents observed in excitable tissues^{14–19}.

We thank Drs Paul Berg and Hiroto Okayama for providing us with their high-efficiency cloning system and Dr Takashi Miyata and Mr Hidenori Hayashida for computer analysis. This investigation was supported in part by research grants from the Ministry of Education, Science and Culture of Japan, the Mitsubishi Foundation and the Japanese Foundation of Metabolism and Diseases.

Received 15 October 1985; accepted 15 January 1986.

- Hodgkin, A. L. & Huxley, A. F. *J. Physiol., Lond.* **117**, 500–544 (1952).
- Agnew, W. S. *A. Rev. Physiol.* **46**, 517–530 (1984).
- Catterall, W. A. *Science* **223**, 653–661 (1984).
- Agnew, W. S., Levinson, S. R., Brabson, J. S. & Raftery, M. A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2606–2610 (1978).
- Miller, J. A., Agnew, W. S. & Levinson, S. R. *Biochemistry* **22**, 462–470 (1983).
- Lombet, A. & Lazdunski, M. *Eur. J. Biochem.* **141**, 651–660 (1984).

- Hartshorne, R. P. & Catterall, W. A. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4620–4624 (1981).
- Barchi, R. L. *J. Neurochem.* **40**, 1377–1385 (1983).
- Noda, M. *et al. Nature* **312**, 121–127 (1984).
- Lombet, A., Renaud, J.-F., Chicheportiche, R. & Lazdunski, M. *Biochemistry* **20**, 1279–1285 (1981).
- Lombet, A., Kazazoglou, T., Delpont, E., Renaud, J.-F. & Lazdunski, M. *Biochem. biophys. Res. Commun.* **110**, 894–901 (1983).
- Renaud, J.-F. *et al. J. Biol. Chem.* **258**, 8799–8805 (1983).
- Sherman, S. J. & Catterall, W. A. in *Regulation and Development of Membrane Transport Processes* (ed. Graves, J. S.) 237–263 (Wiley, New York, 1985).
- Barrett, J. N. & Crill, W. E. *J. Physiol., Lond.* **304**, 231–249 (1980).
- Gundersen, C. B., Miledi, R. & Parker, I. *Proc. R. Soc. B220*, 131–140 (1983).
- Jaimovich, E. *et al. Eur. J. Physiol.* **397**, 1–5 (1983).
- Eick, R. T., Yeh, J. & Matsuki, N. *Biophys. J.* **45**, 70–73 (1984).
- Gilly, W. F. & Armstrong, C. M. *Nature* **309**, 448–450 (1984).
- Benoit, E., Corbier, A. & Dubois, J.-M. *J. Physiol., Lond.* **361**, 339–360 (1985).
- Costa, M. R. C., Casnellie, J. E. & Catterall, W. A. *J. Biol. Chem.* **257**, 7918–7921 (1982).
- Costa, M. R. C. & Catterall, W. A. *J. Biol. Chem.* **259**, 8210–8218 (1984).
- Krebs, E. G. & Beavo, J. A. *Rev. Biochem.* **48**, 923–959 (1979).
- Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105–132 (1982).
- Chou, P. Y. & Fasman, G. D. *A. Rev. Biochem.* **47**, 251–276 (1978).
- Hille, B. *Ionic Channels of Excitable Membranes* (Sinauer, Sunderland, Massachusetts, 1984).
- Armstrong, C. M. & Bezanilla, F. *Nature* **242**, 459–461 (1973).
- Capaldi, R. A. & Vanderkooi, G. *Proc. natn. Acad. Sci. U.S.A.* **69**, 930–932 (1972).
- Engelman, D. M. & Steitz, T. A. *Cell* **23**, 411–422 (1981).
- Rice, C. M., Bell, J. R., Hunkapiller, M. W., Strauss, E. G. & Strauss, J. H. *J. molec. Biol.* **154**, 355–378 (1982).
- Hubbard, S. C. & Ivatt, R. J. *A. Rev. Biochem.* **50**, 555–583 (1981).
- Bause, E. *Biochem. J.* **209**, 331–336 (1983).
- Armstrong, C. M. *Physiol. Rev.* **61**, 644–683 (1981).
- Hille, B. *J. gen. Physiol.* **58**, 599–619 (1971).
- Hille, B. *J. gen. Physiol.* **59**, 637–658 (1972).
- Dayhoff, M. O., Schwartz, R. M. & Orcutt, B. C. in *Atlas of Protein Sequence and Structure* Vol. 5, Suppl. 3 (ed. Dayhoff, M. O.) 345–352 (National Biomedical Research Foundation, Silver Spring, Maryland, 1978).
- Okayama, H. & Berg, P. *Molec. cell. Biol.* **2**, 161–170 (1982).
- Proudfoot, N. J. & Brownlee, G. G. *Nature* **263**, 211–214 (1976).
- Goeddel, D. V. *et al. Nature* **290**, 20–26 (1981).
- Berk, A. J. & Sharp, P. A. *Cell* **12**, 721–732 (1977).
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294–5299 (1979).
- Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1408–1412 (1972).
- Noda, M. *et al. Nature* **295**, 202–206 (1982).
- Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201–5205 (1980).
- Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).
- Payvar, F. & Schimke, R. T. *J. Biol. Chem.* **254**, 7636–7642 (1979).
- Weinstock, R., Sweet, R., Weiss, M., Cedar, H. & Axel, R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1299–1303 (1978).

Multiple conformations of a protein demonstrated by magnetization transfer NMR spectroscopy

Robert O. Fox*, Philip A. Evans & Christopher M. Dobson

Inorganic Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QR, UK

It is generally accepted that a globular protein in its native state adopts a single, well-defined conformation. However, there have been several reports that some proteins may exist in more than one distinct folded form in equilibrium. In the case of staphylococcal nuclease, evidence for multiple conformations has come from electrophoretic and NMR studies^{1–4}, although there has been some controversy as to whether these are actually interconvertible forms of the same molecular species^{5,6}. Recently, magnetization transfer (MT)-NMR has been developed as a means of studying the kinetics of conformational transitions in proteins⁷. In the study reported here, this approach has been extended and used to demonstrate the presence of at least two native forms of nuclease in equilibrium and to study their interconversion with the unfolded state under the conditions of the thermal unfolding transition. The experiments reveal that two distinct native forms of the protein fold and unfold independently and that these can interconvert directly as well as via the unfolded state. The spectra of the different forms suggest that they are structurally similar but the MT experiments show that the kinetics of folding and unfolding are quite

* Present address: Department of Cell Biology, Stanford University Medical School, Stanford, California 94305, USA.

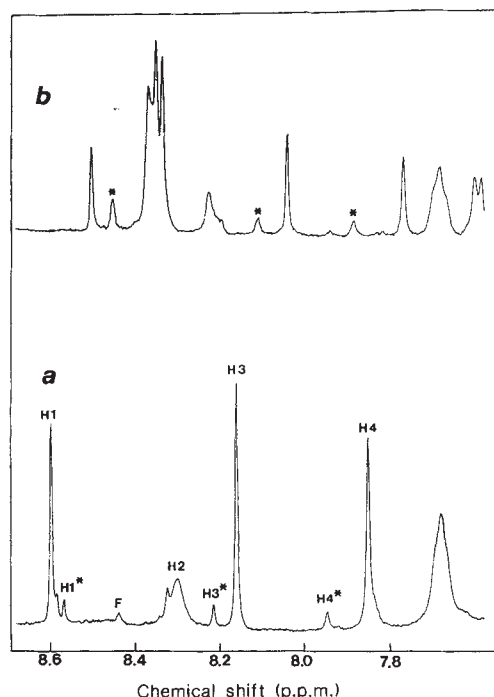


Fig. 1 *a*, 500-MHz ^1H -NMR spectrum of nuclease (100 mg ml^{-1}), pH 5.5 in 200 mM deuterated sodium acetate solution (37°C). Backbone amide protons have been exchanged for deuterons to reveal the resonances of the His H^ϵ protons. The asterisks denote histidine resonances of minor folded forms of the protein. F denotes a non-protein impurity. *b*, Spectrum of nuclease; same sample as in *a* but at 55°C .

Methods. The nuclease used in these experiments was produced by a recombinant expression construct in which the nuclease gene cloned from *S. aureus* Foggi strain was under the control of the λ_{PL} promoter in the vector pAS1 (R.O.F., unpublished). The protein produced contains the additional N-terminal sequence fMet-Asp-Pro-Thr-Val-Tyr-Ser beyond the nuclease A sequence¹⁹. Nuclease produced in *E. coli* was isolated to >95% purity on a phosphocellulose column developed with a linear gradient of 0.3 M ammonium acetate, pH 6 to 1.0 M ammonium acetate, pH 8 (ref. 20).

different. Characterization of this behaviour will, therefore, have important implications for our understanding of the relationship between structure and folding kinetics.

The Ca^{2+} -dependent nuclease of *Staphylococcus aureus* (EC 3.1.4.7) has been used extensively in folding studies because of its small size (149 residues), the reversibility of unfolding *in vitro* and the absence of disulphide crosslinks (reviewed in refs 8–10). The structure of a crystal of the nuclease has been determined to 1.5-Å resolution, revealing extensive α - and β -structure in the folded state^{11,12}. In the present study we have used a protein obtained by expression of the recombinant nuclease gene in *Escherichia coli*. This protein includes additional residues at the N-terminus, part of a sequence which is cleaved from the mature protein in *S. aureus*.

Part of the ^1H -NMR spectrum of the nuclease in native conditions is shown in Fig. 1*a*. The H^ϵ (C2H) resonances of the four histidines are prominent and are accompanied by several smaller peaks. Each minor resonance has a chemical shift comparable to that of one of the major peaks and these relationships persist throughout the pH range over which the resonances titrate. This result suggests that the minor peaks are due to one or more folded forms which are native-like but distinct from the major form. Here we consider one pair of resonances (labelled H4 and H4* in Fig. 1) which are sufficiently well resolved to permit detailed NMR studies.

The relative intensities of the major (H4) and minor (H4*) resonances change reversibly with temperature, which shows

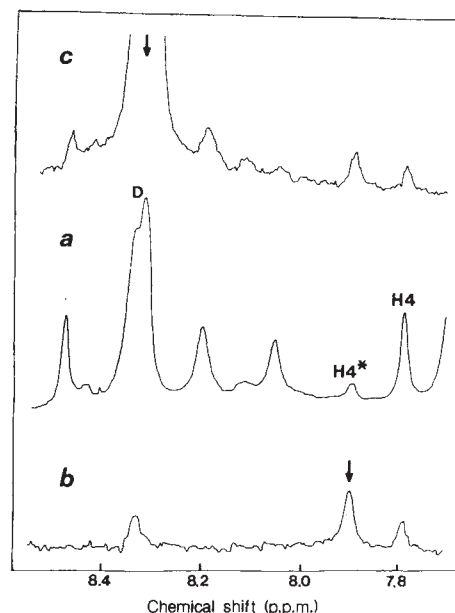


Fig. 2 Single-saturation magnetization transfer experiments on nuclease, pH 5.5 (55°C). *a*, A 'blank' spectrum in which a pre-saturation pulse was applied at a frequency at which no protons resonate. *b*, *c*, Difference spectra obtained by subtraction from *a* of spectra acquired in an interleaved fashion in which individual resonances (denoted by the arrows) have been selectively saturated. The vertical scale of the difference spectra is four times that of spectrum *a*. In *b*, H4*, a resonance of a minor folded state, was irradiated for 3 s before observation of the spectrum. In *c* the saturation pulse was applied to the resonances of the unfolded state. The selectivity of the saturation in these experiments was checked by carrying out control experiments at lower temperatures where there is no measurable magnetization transfer.

that they arise from different states of the same molecule in equilibrium. Moreover, on forming a ternary complex with Ca^{2+} and thymidine 3',5'-diphosphate, H4* is lost from the nuclease spectrum, suggesting that the stabilization afforded by binding the ligands displaces the equilibrium in favour of the major folded form.

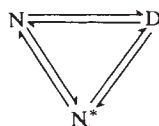
In the spectrum shown in Fig. 1*b*, the temperature has been raised to 55°C , which is in the transition zone of reversible thermal unfolding. The resonances of the folded forms are now accompanied by those of the unfolded protein, whose His H^ϵ peaks are clustered between 8.3 and 8.4 parts per 10^6 (p.p.m.). We have studied the processes of interconversion of the different states through a series of MT experiments. If a resonance of the protein in one state is selectively perturbed, the corresponding resonances in the other states may in turn be affected as a result of chemical exchange between the states. Figure 2 presents difference spectra showing the effects of saturating individual resonances and allowing the system to equilibrate in the presence of the perturbation. In these conditions the resonance intensity of a site in exchange with that which has been irradiated will depend on the rate of magnetization transfer from this site compared with its spin-lattice relaxation rate^{7,13,14}.

Figure 2*b* shows the effect of selective saturation of H4* in the conditions of the thermal unfolding transition. Decreases in the intensities of H4 and of the corresponding unfolded-state resonance, D, demonstrate directly that all three states are interconverting under these conditions. Figure 2*c* shows the effect of saturating the D resonance: as expected, both H4 and H4* are affected but the effect on H4* is proportionately much greater. As the spin-lattice relaxation times of H4 and H4* are comparable, this result shows that the equilibrium unfolding rate of the minor species, N*, is much greater than that of the major folded form, N. The effect on D of saturating H4 is also much smaller than when H4* is saturated, which shows that the overall rate of conversion of D to N is smaller than that to N*,

although further interpretation of this result is complicated by the possibility that the unfolded state is not kinetically homogeneous. We shall discuss this further below.

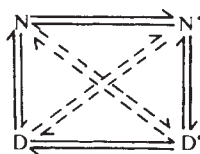
These single-saturation experiments alone do not reveal the pathways of interconversion in such a multi-site system. To address this problem we have extended the MT approach to measure the effects of simultaneous saturation of two resonances. Figure 3 compares the effect of saturating H4* alone with saturating H4* and D together. The difference spectrum shows clearly that the extent of saturation transfer to H4 increased when both resonances were saturated; this means that magnetization transfer can occur directly between N and D without the need to be relayed through N*, excluding the possibility that N* is a required intermediate for the folding and unfolding of N. Similarly, saturation of both D and H4 leads to a larger effect on H4* than saturation of H4 alone, which reveals direct magnetization transfer from H4* to D and thus confirms that the different native forms unfold independently. Furthermore, we found that while D was saturated, direct magnetization transfer between H4 and H4* was detectable; this shows that the different folded forms can interconvert without the need to pass through the unfolded state. Magnetization transfer between H4 and H4* was not measurable at significantly lower temperatures, however, showing that there is a substantial energy barrier to this process.

The above experiments have thus revealed mutual exchange between at least three states of nuclease under the experimental conditions:



This conclusion is based on the behaviour of one of the four histidines, H4. Minor resonances are also evident for the other histidines and elsewhere in the spectrum. At least some of these resonances show behaviour very similar to H4* in a two-dimensional MT experiment and it seems likely that these arise from a common minor component. It has been possible, however, to observe that the intensity of another minor peak changes relative to that of H4* on variation of the temperature. Further, unlike H4*, it is still present in the spectrum of the inhibited nuclease. This result shows that there could, in fact, be more than a single equilibrium among folded species. Although it is not possible to define the conformations of the different states, they are clearly not very different. For example, each of the minor histidine resonances has a chemical shift similar to that of one of the major peaks, and their pH dependences demonstrate that the pK_a values in the different states are very similar.

Another aspect which must be considered concerns the possibility of spectroscopically unresolved heterogeneity of the unfolded state arising from *cis-trans* isomerism of proline peptide bonds¹⁵, which is of particular interest because this could also be the origin of the heterogeneity of the folded protein. It is known that RNase A, for example, has a metastable form which contains a non-native proline isomer¹⁶, and it is possible that in some circumstances such species could be sufficiently stable to be populated at equilibrium. In this case the scheme would be extended thus:



The MT experiment will allow us to test these models by examining the time course of development of the effects.

Globular proteins have usually been observed to equilibrate rapidly about a single average structure in solution. Evidence

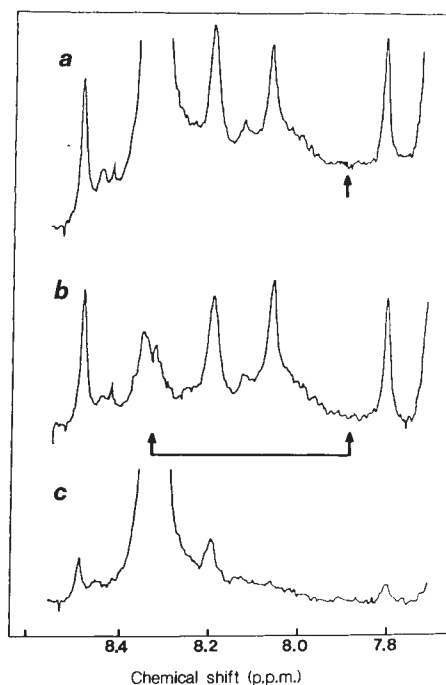


Fig. 3 Double-saturation magnetization transfer experiment on nuclease, pH 5.5 (55 °C). In spectrum *a* the H4* resonance was saturated for 3 s before observation. In spectrum *b*, which was acquired in an interleaved fashion, both H4* and the resonances of the unfolded state were saturated. *c*, The differences between the two spectra. The simultaneous double saturation was achieved by applying a single selective 3-s pulse to H4* in the normal way while simultaneously applying a DANTE pulse²¹ of the same duration to the resonances of the unfolded state.

now exists, however, for multiple folded forms of similar stability in the cases of nuclease and some other proteins, including dihydrofolate reductase^{17,18}. We have demonstrated here that MT studies constitute a powerful tool for the elucidation of pathways of conformational interconversion in these systems. We anticipate that the existence of alternative stable folded states will provide unique possibilities for elucidating the effects of single-site mutations on folding.

This work was supported by SERC. R.O.F. acknowledges a postdoctoral fellowship from the Jane Coffin Childs Memorial Fund for Medical Research whilst in Oxford and the encouragement and support of N. D. F. Grindley and his laboratory in carrying out the cloning and expression of the nuclease gene. C.M.D. is a member of the Oxford Enzyme Group.

Received 11 December 1985; accepted 14 January 1986.

1. Taniuchi, H. & Anfinsen, C. B. *J. biol. Chem.* **241**, 4366-4385 (1966).
2. Markley, J. L., Williams, M. N. & Jardetzky, O. *Proc. natn. Acad. Sci. U.S.A.* **65**, 645-651 (1970).
3. Arata, Y., Khalifah, R. & Jardetzky, O. *Ann. N.Y. Acad. Sci.* **222**, 230-239 (1973).
4. Tucker, P. W., Hazen, E. E. & Cotton, F. A. *Molec. cell. Biochem.* **22**, 67-77 (1978).
5. Cohen, J. S., Shrager, R. I., McNeill, M. & Schechter, A. N. *Nature* **228**, 642-644 (1970).
6. Marrakechi, N. & Cozzone, P. J. *C. r. heb. Séanc. Acad. Sci., Paris Ser. 3* **293**, 23-26 (1981).
7. Dobson, C. M. & Evans, P. A. *Biochemistry* **23**, 4267-4270 (1984).
8. Anfinsen, C. B., Schechter, A. N. & Taniuchi, H. *Cold Spring Harb. Symp. quant. Biol.* **36**, 243-249 (1971).
9. Anfinsen, C. B. *Science* **181**, 223-230 (1973).
10. Tucker, P. W., Hazen, E. E. & Cotton, F. A. *Molec. cell. Biochem.* **23**, 131-141 (1979).
11. Arnone, A. *et al. J. biol. Chem.* **246**, 2302-2316 (1971).
12. Cotton, F. A., Hazen, E. E. & Legg, M. J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2551-2555 (1979).
13. Forsén, S. & Hoffman, R. A. *J. chem. Phys.* **39**, 2892-2901 (1963).
14. Campbell, I. D., Dobson, C. M., Ratcliffe, R. G. & Williams, R. J. P. *J. magn. Reson.* **29**, 397-417 (1978).
15. Brandts, J. F., Halvorson, H. R. & Brennan, M. *Biochemistry* **14**, 4953-4963 (1975).
16. Cook, K. H., Schmid, F. X. & Baldwin, R. L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6157-6161 (1979).
17. Gronenborn, A. *et al. Molec. Pharm.* **20**, 145-153 (1981).
18. Matthews, C. R., Perry, K. M. & Touchette, N. A. *Am. chem. Soc. Abstr.* **186**, BIOL83 (1983).
19. Cone, J. L., Cusumano, C. L., Taniuchi, H. & Anfinsen, C. B. *J. biol. Chem.* **246**, 3103-3110 (1971).
20. Fuchs, S., Cuatrecasas, P. & Anfinsen, C. B. *J. biol. Chem.* **242**, 4768-4770 (1967).
21. Morris, G. A. & Freeman, R. *J. magn. Reson.* **29**, 433-462 (1978).

Predisposition to helminth infection in man

I FIND it difficult to accept the basic assumptions made by Anderson and May¹ in their model of herd immunity, at least as it applies to schistosome infection. They initially assume that certain individuals are predisposed to heavy infections. In field studies of individuals with treated *Schistosoma mansoni* infections³, this was not the case. The supposition is supported by studies of hookworm infections⁴. Even if predisposition to heavy schistosome infection were shown to exist, it would be premature to attribute this to lack of immunity or to assume that such individuals cannot be successfully vaccinated.

Anderson and May also assume that immunity to schistosome infection is correlated with the intensity of infection and cite the more rapid decline in faecal egg counts in a heavily infected population examined by Siongok *et al.*⁵ compared with more lightly infected persons examined by Abdel-Wahab *et al.*⁶. My objections here are that the difference in faecal egg counts between these populations was not great (464 eggs per g of faeces in 5-9-yr-old males in the study by Abdel-Wahab *et al.*⁶ and 650 eggs per g in the groups examined by Siongok *et al.*⁵) and the report of Ongom and Bradley⁷ is not cited. In this last study the decline in egg excretion was gradual although 5-9-yr-old boys excreted an average of 793 eggs per g faeces. In the very lightly infected populations in Puerto Rico and St Lucia (geometric means of 20-30 eggs per g faeces in 5-9-yr-olds), egg excretion was maintained at relatively stable levels for many years^{8,9}. Resistance of mice to reinfection is clearly related to infection intensity¹⁰, but a large part of this resistance is nonspecific¹¹ and is related to infection intensities that generally are not relevant for humans¹².

The experimental study by Crombie and Anderson¹³ also presents problems and does not provide convincing evidence for the death of significant numbers of *S. mansoni* worms in mice. Although dead worms are visible in the livers of mice for months after treatment¹⁴, I rarely encounter them in the livers of untreated mice. Certainly, the more heavily infected mice die sooner than the less heavily infected ones, a point considered by the authors but dismissed, without presentation of data, as a cause of decreasing intensity of infection with time. The mouse model is an inherently difficult one in which to pursue this point, as nearly all chronically infected mice develop large portal-systemic collateral veins with resultant shunting of some worms into the pulmonary circulation (my unpublished data), where their survival is probably

limited. Such collaterals also develop in patients with severe disease, but such patients form a small part of the population and would not be important epidemiologically.

ALLEN W. CHEEVER

Laboratory of Parasitic Diseases,
National Institute of Allergy
and Infectious Diseases,
National Institutes of Health,
Bethesda, Maryland 20205, USA

1. Anderson, R. M. & May, R. M. *Nature* **315**, 493-496 (1985).
2. Mahmoud, A. A. F., Siongok, T. A., Ouma, J., Houser, H. B. & Warren, K. S. *Lancet* **i**, 849-851 (1983).
3. Sturrock, R. F. *et al. Trans. R. Soc. trop. Med. Hyg.* **77**, 363-371 (1983).
4. Shad, G. A. & Anderson, R. M. *Science* **228**, 1537-1540 (1985).
5. Siongok, T. K. A. *et al. Am. J. trop. Med. Hyg.* **25**, 273-284 (1976).
6. Abdel-Wahab, M. F. *et al. Am. J. trop. Med. Hyg.* **29**, 868-874 (1980).
7. Ongom, V. L. & Bradley, D. J. *Trans. R. Soc. trop. Med. Hyg.* **66**, 835-851 (1972).
8. Hiatt, R. A., Cline, B. L., Ruiz-Tiben, E., Knight, W. B. & Berrios-Duran, L. A. *Am. J. trop. Med. Hyg.* **29**, 1228-1240 (1980).
9. Barnish, G., Jordan, P., Bartholomew, R. K. & Grist, E. *Trans. R. Soc. trop. Med. Hyg.* **76**, 602-609 (1982).
10. Dean, D. A. *Expl. Parasit.* **55**, 1-104 (1983).
11. Dean, D. A., Minard, P., Murrell, K. D. & Vannier, W. E. *Am. J. trop. Med. Hyg.* **27**, 957-965 (1978).
12. Cheever, A. W. *Trans. R. Soc. trop. Med. Hyg.* **63**, 781-795 (1969).
13. Cheever, A. W., DeWitt, W. B. & Warren, K. S. *Am. J. trop. Med. Hyg.* **14**, 239-253 (1965).

ANDERSON *ET AL.* REPLY—The assumption of predisposition to heavy (or light) infection is not central to the framework of our model¹ of herd immunity to helminth infection (see equations (1) and (2) of ref. 1). It is simply an optional extension which can be used to investigate the potential influence of heterogeneity in either exposure to infection (that is, behavioural factors) or the host's ability to restrict parasite establishment and survival (that is, immunological factors which may be under genetic control) on the transmission and control of helminths in human communities. Such heterogeneity is a central feature of the epidemiology of helminths, as indicated by the high degrees of aggregation observed in parasite distributions within human populations². However, the interesting issue raised by Cheever centres on the question of whether or not predisposition is an important component of the epidemiology of schistosome parasites. We believe it is. The two studies cited by Cheever do not address the statistical question of whether there is a positive association (in large samples of people, stratified according to age and sex) between faecal egg counts in individual patients before treatment and after an interval of reinfection following chemotherapy^{3,4}. The first such analysis of this question⁷ for *Schistosoma mansoni* infection in a rural community in Kenya^{5,6}, reveals a highly significant posi-

tive association (12-month period of reinfection in 8-15-yr-old children, Kendall's $\tau = 0.347$, $n = 117$, $P < 0.0001$; ref. 7) between pretreatment egg counts (in eggs per g, e.p.g.) and counts following an interval of reinfection after chemotherapeutic treatment⁷. Whether this observed pattern is a consequence of differences in contact with infection, or in immunological competence of the host, is unclear at present. Work in progress (the same field study^{5,6}) on individual patterns of contact with water and faecal egg output should improve our understanding of this issue. More broadly, recent field studies record evidence for predisposition to heavy infection with *Ascaris*^{8,9}, hookworm¹⁰ and *Trichuris*¹¹.

We disagree with Cheever's interpretation of the relevant data on the question of a positive association between the force of transmission and the degree of convexity of age-intensity of infection profiles (Fig. 1b in ref. 1). In the 5-9-yr-old age classes in the two studies we cited^{12,13}, the intensity of infection (mean e.p.g. per person sampled, including boys and girls) was 690 eggs per g faeces in the high transmission area and 144 eggs per g in the low transmission area. Furthermore, in the study of Ongom and Bradley¹⁴ (within a high transmission area of Uganda), the average intensity of infection (boys and girls) fell from a maximum of 1,128 eggs per g of faeces in the 5-9-yr-old class to a minimum of 404 eggs per g of faeces in the 40-49-yr-old class; again, a marked convex pattern was associated with a high net force of transmission. The studies in Puerto Rico and St. Lucia^{15,16}, cited by Cheever, further support our argument, as in these areas of low transmission intensity, mean faecal egg counts remain relatively stable over a wide range of age classes. More generally, a recent statistical study, using a range of epidemiology surveys for *S. mansoni* and *Schistosoma haematobium*, of the relationship between the maximum average intensity of infection (commonly in the child/teenage age groups) and the rate of decline in average intensity from child to adult groups reveals a highly significant positive correlation (J.A.C., unpublished).

With respect to our experimental studies¹⁷, juvenile and dead adult worms were recovered from mice throughout the long periods of repeated exposure to infection; this suggests that populations of adult parasites in individual mice are subject to continual recruitment and mortality. We discounted parasite-induced host mortality as the major cause of the convex profiles of change in mean parasite burden with duration of exposure (= mouse age) (Fig. 1a in ref. 17) as at all sampling points in the trickle studies, the distribution of worm numbers per mouse

in each sample remained Poisson in form (as predicted by stochastic models of immigration-death processes¹⁸). Within a cohort of inbred mice repeatedly exposed to the same level of infection, underdispersion in parasite numbers per host would have resulted, especially in the latter part of the experiment when mean worm burdens declined, if the host mortality were linked directly to worm burden¹⁹ (the rapid elimination of heavily infected animals acting to decrease the variance within a sample).

More generally, we agree that mice are not ideal models for the study of schistosome infections of man. However, our aim was to show that experimental studies of immunity to helminth parasites, based on repeated exposure to low levels of infection (as is thought to occur in human communities within areas of endemic infection), produce results that are not always in accord with more conventional experimental designs based on primary infection and a single subsequent challenge.

R. M. ANDERSON
J. A. CROMBIE

Department of Pure and
Applied Biology,
Imperial College,
University of London,
London SW7 2BB, UK

R. M. MAY

Department of Biology,
University of Princeton,
Princeton,
New Jersey 08540, USA

Cause of the 'inhibitor' phenotype in the haemophilias

Following the paper by Gitschier *et al.*¹, we feel that a brief comment on their data and our earlier findings in haemophilia B² may help to clarify, or generate ideas that may lead to a better understanding of, the causes of the 'inhibitor' phenotype, a necessary step in the development of a completely successful treatment.

In 1983 we reported gross gene deletions in four out of five haemophilia B patients with the inhibitor phenotype, in support of the hypothesis that haemophilia B patients with inhibitors are predisposed to develop antibodies to factor IX as a result of mutations that prohibit the synthesis of a sufficiently normal factor IX protein and thus prevent the development of immune tolerance to normal factor IX. Of course, whether such a predisposition results in the inhibitor complication may depend on other factors such as the patient's genetic background and environmental experience: for example, the intensity and duration of treatment with normal coagulants. In four of our patients the predisposing mutation was a deletion of two-thirds or more of the factor IX coding sequence, but, of course, point mutations, frameshift mutations or shorter deletions could have had the same effect if they had occurred at positions which prevented the synthesis of a protein capable of inducing immune tolerance to normal factor IX. For example, the predisposing mutations might interfere with synthesis or processing of RNA, or might be mutations that affect the translation of messenger RNA into protein.

Previously, we predicted that gross changes of the factor VIII(c) gene might be found in haemophilia A patients with inhibitors. Now, Gitschier *et al.*¹ report a patient with a deletion involving exon 26 which codes for the last 51 amino acids of factor VIII, and another with an aberrant translation stop codon arresting translation of the factor VIII mRNA 26 amino acids prematurely, neither of which are associated with antibodies to factor VIII. Conversely, a patient with a different aberrant stop codon generating a truncated protein that is 124 amino acids short and another patient with a deletion comprising exons 23–25 which code for 157 amino acids, have both developed inhibitors. Such results seem in keeping with our observations in haemophilia B, since the loss of normal factor VIII epitopes in the latter two cases would be expected to be greater than in the first two cases of Gitschier *et al.*¹, and hence more likely to result in defective immune tolerance to normal factor VIII.

In haemophilia B four further inhibitor patients have shown gross gene deletions: one in the United Kingdom³ and three in Italy (H. J. Hassan *et al.*, F. Bernardi *et*

al. and M. Pecorara *et al.*, personal communications). These findings, of course, are at variance with the results of Gitschier *et al.* in haemophilia A, as many of their inhibitor patients did not show easily detectable gene defects. However, this is not the crux of the matter. What matters is whether the inhibitor complication arises in patients with mutations preventing the development of immune tolerance or not. In other words, is the discrepancy between the findings in haemophilia A and B principally a result of a more complex aetiology of the inhibitor complication in haemophilia A or caused by the plurality of factor VIII defects which can prevent the development of immune tolerance to normal factor VIII, in keeping with the example of haemophilia B?

We believe that the characterization of both the gene defects and the epitope specificity of the antibodies against factor VIII found in inhibitor patients would clarify the situation and reveal important immunological features of this coagulation factor.

F. GIANNELLI

Paediatric Research Unit,
United Medical Schools of
Guy's and St Thomas's Hospitals,
Guy's Tower, London Bridge,
London SE1 9RT, UK

G. G. BROWNLEE

Sir William Dunn School of Pathology,
South Parks Road,
Oxford OX1 5RE, UK

1. Gitschier, J. *et al.* *Nature* **315**, 427–430 (1985).
2. Giannelli, F. *et al.* *Nature* **303**, 181–182 (1983).
3. Peake, I. R., Furlong, B. L. & Bloom, A. L. *Lancet* **i**, 242–243 (1984).

1. Anderson, R. M. & May, R. M. *Nature* **315**, 493–496 (1985).
2. Anderson, R. M. & May, R. M. *Adv. Parasit.* **25**, 1–101 (1985).
3. Mahmoud, A. A. F., Arap Siongok, T. A., Ouma, J., Houser, H. B. & Warren, K. S. *Lancet* **i**, 849–851 (1983).
4. Wilkins, H. A., Goll, P. H., Marshall, T. F. de C. & Moore, P. J. *Trans. R. Soc. trop. Med. Hyg.* **78**, 216–221 (1984).
5. Butterworth, A. E. *et al.* *Trans. R. Soc. trop. Med. Hyg.* **78**, 108–123 (1984).
6. Butterworth, A. E. *et al.* *Trans. R. Soc. trop. Med. Hyg.* **79**, 393–408 (1985).
7. Bensted-Smith, R. *et al.* *Trans. R. Soc. trop. Med. Hyg.* (in the press).
8. Elkins, D., Haswell-Elkins, M. & Anderson, R. M. *Trans. R. Soc. trop. Med. Hyg.* (in the press).
9. Hlilang, T. in *Ascariasis and its Public Health Significance* (eds Crompton, D. W. T., Nesheim, M. C. & Pawlowski, Z.) 21–48 (Taylor & Francis, London, 1985).
10. Schad, G. A. & Anderson, R. M. *Science* **228**, 1527–1540 (1985).
11. Bundy, D. A. P. *Trans. R. Soc. trop. Med. Hyg.* (in the press).
12. Arap Siongok, T. K. *et al.* *Am. J. trop. Med. Hyg.* **25**, 274–284 (1976).
13. Abdel-Wahab, M. F. *et al.* *Am. J. trop. Med. Hyg.* **28**, 868–874 (1980).
14. Ongom, V. L. & Bradley, D. J. *Trans. R. Soc. trop. Med. Hyg.* **66**, 835–851 (1972).
15. Hailt, R. A., Cline, B. L., Ruiz-Tiben, E., Knight, W. B. & Berrois-Duran, L. A. *Am. J. trop. Med. Hyg.* **29**, 1228–1240 (1980).
16. Barnish, G., Jordon, P., Bartholomew, R. K. & Crist, E. *Trans. R. Soc. trop. Med. Hyg.* **76**, 602–609 (1982).
17. Crombie, J. A. & Anderson, R. M. *Nature* **315**, 419–483 (1985).
18. Anderson, R. M. & Medley, G. *Parasitology* **90**, 629–660 (1985).
19. Anderson, R. M. & Gordon, R. *Parasitology* **85**, 373–398 (1982).

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):

- ☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnType ☐ OCLC ILL Subsystem
☐ Other (please specify) _____
☐ I am interested in sending my order by mail.
☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name _____
 Title _____
 Institution/Company _____
 Department _____
 Address _____
 City _____ State _____ Zip _____
 Phone (____) _____

UMI Article
Clearinghouse

Mail to: University Microfilms International
 300 North Zeeb Road, Box 91 Ann Arbor, MI 48106

Mathematics

Calculus. By HARLEY FLANDERS. W.H. Freeman: 1985. Pp. 984. ISBN 0-7167-1643-7. Np.

Graph Theory With Applications to Algorithms and Computer Science. Y. ALVI, G. CHARTRAND, L. LESNIAK, D.R. LICK and G.E. WALL (eds). Wiley: 1985. Pp. 810. Hbk ISBN 0-471-81635-3. £46.

Integration Theory. Vol. 1. Measure and Integral. By CORNELIU CONSTANTINESCU and KARL WEBER. Wiley: 1985. Pp. 520. ISBN 0-471-04479-2. £50.65.

Oxford Logic Guides, Vol. 12: Boolean-Valued Models and Independence Proofs in Set Theory, 2nd Edn. By J.L. BELL. Clarendon Press: 1985. Pp. 165. ISBN 0-19-853241-5. £16.

Principia Mathematica. By ISAAC NEWTON. Christian Bourgeois Editeur: 1985. Pp. 376. Pbk ISBN 2-267-00425-9. Np.

Astronomy

Anuario del Observatorio Astronómico 1985. INSTITUTO GEOGRÁFICO NACIONAL. Talleres del Instituto Geográfico, Madrid: 1985. Pp. 502. ISBN 84-505-1088-0. Pts 400.

Astronomy and Astrophysics Abstracts, Vol. 38. S. BOHME *et al.* (eds). Springer-Verlag: 1984. Pp. 919. ISBN 3-540-15562-7/0-387-15562-7. DM 189.

Astronomy and Astrophysics Publications from D. Reidel Publishing Company, Fall 1985. D. Reidel: 1985. Pp. 200. Pbk np.

The Atmosphere of Venus: Recent Findings. G.M. KEATING, A.J. KLIORE and V.I. MOROZ (eds). Pergamon: 1985. Pp. 201. Pbk ISBN 0-08-033223-4. Np.

Birth and Infancy of the Stars. Proceedings of the Les Houches Summer School, Session XLI, 8 August–2 September, 1983. Elsevier: 1985. Pp. 846. Hbk ISBN 0-444-86917-4. Dfl. 430. \$159.25.

Glimpsing an Invisible Universe: The Emergence of X-Ray Astronomy. By RICHARD F. HIRSH. Cambridge University Press: 1985. Pp. 186. Pbk ISBN 0-521-31232-9. Pbk £8.95.

Ices in the Solar System. JÜRGEN KLINGER *et al.* (eds). Reidel: 1985. Pp. 954. ISBN 90-277-2062-2. Dfl. 290, \$99, £80.50.

Planetary Landscapes. By R. GREELEY. Allen & Unwin: 1985. Pp. 265. Hbk ISBN 0-04-551080-6. £44.95.

Protostars and Planets II. DAVID C. BLACK and MILDRED SHAPLEY MATTHEWS (eds). University of Arizona Press: 1985. Pp. 1293. Hbk ISBN 0-8165-0950-6. \$45.

Venus: An Errant Twin. By ERIC BURGESS. Columbia University Press: 1985. Pp. 160. Hbk ISBN 0-231-05856-X. \$29.95.

Physics

Advances in Infrared and Raman Spectroscopy, Vol. 12. R.J.H. CLARK and R.E. HESTER (eds). Wiley: 1985. Pp. 360. ISBN 0-471-90674-3. £72.

Applications of Plasma Processes to VLSI Technology. TAKUO SUGANO (ed.). Wiley: 1985. Pp. 394. Hbk ISBN 0-471-86960-0. £46.

Applied Classical Electrodynamics, Vol. 1: Linear Optics. F.A. HOPF and G.I. STEGEMAN. Wiley: 1985. Pp. 262. Hbk ISBN 0-471-82788-6. £30.65.

Classical Electrodynamics. By ROMAN S. INGARDEN and ANDREZJ JAMIOŁKOWSKI. Elsevier: 1985. Pp. 350. ISBN 0-444-99604-4. \$86.50.

Computer Aided Chemical Thermodynamics of Gases and Liquids: Theory, Models, Programs. By PAUL BENEDEK and FERENC OLT. Wiley: 1985. Pp. 731. Hbk ISBN 0-471-87825-1. £86.95.

Current Topics in Photovoltaics. T.J. COUTTS and J.D. MEAKIN (eds). Academic: 1985. Pp. 279. ISBN 0-12-193860-3. \$62, £50.

Diffraction Processes in Nuclear Physics. By W.E. FRAHN. Oxford University Press: 1985. Pp. 179. ISBN 0-19-851512-X. £22.50.

Galaxies, Axisymmetric Systems and Relativity. M.A.H. MACCALLUM (ed.). Cambridge University Press: 1985. Pp. 300. Hbk ISBN 0-521-30812-7. £25, \$49.50.

High-energy Ion-atom Collisions, Second Workshop. D. BERÉNYI and G. HOCK (eds). Akadémiai Kiadó, Budapest: 1985. Pp. 306. ISBN 963-05-4091-6. £28.50.

Integral Systems in Statistical Mechanics. G.M. D'ARIANO, A. MONTORSI and M.G. RASETTI (eds). World Scientific: 1985. Pp. 169. ISBN 9971-978-11-3. £29.65.

Introduction to Magnetic Recording. ROBERT M. WHITE (ed.). Institute of Electrical and Electronics Engineers: 1985. Pp. 307. ISBN 0-87942-184-3. Np.

Memory Function Approaches to Stochastic Problems in Condensed Matter. M.W. EVANS, P. GRIGOLINI and G. PASTORI PARRAVICINI (eds). Wiley: 1985. Pp. 556. Hbk ISBN 0-471-80482-7. £86.90.

The Metallic and Nonmetallic States of Matter. P.P. EDWARDS and C.N.R. RAO (eds). Taylor & Francis: 1985. Pp. 427. Hbk ISBN 0-85066-321-0. £40.

Multiple Photon Infrared Laser Photophysics and Photochemistry. By V.N. BAGRATASHVILI, V.S. LETOKHOV, A.A. MAKAROV and E.A. RYABOV. Harwood Academic Publishers: 1985. Pp. 512. Hbk ISBN 3-7186-0296-5. \$65.

Nonequilibrium Phonon Dynamics. WALTER E. BRON (ed.). Plenum: 1985. Pp. 679. ISBN 0-306-42008-2. \$97.50.

Perspectives in Particles and Fields, Cargèse 1983. MAURICE LEVY, JEAN LOUIS BASDEVANT, DAVID SPEISER, JACQUES WEYERS, MAURICE JACO and RAYMOND GASTMANS (eds). Plenum: 1985. Pp. 598. Hbk ISBN 0-306-42067-8. Np.

Phenomena Induced by Intermolecular Interactions. Nato ASI Series. G. BIRNBAUM (ed.). Plenum: 1985. Pp. 792. Hbk ISBN 0-306-42071-6. Np.

Physics of Disordered Materials. DAVID ADLER, HELLMUT FRITZSCHE and STANFORD R. OVSHINKY (eds). Plenum: 1985. Pp. 850. ISBN 0-306-42074-0. Np.

Problems and Solutions in Electromagnetic Theory. By C.M. LERNER. Wiley: 1985. Pp. 614. Pbk ISBN 0-471-88678-5. £35.75.

Rates of Phase Transformations. By R.H. DOREMUS. Academic: 1985. Pp. 176. Hbk ISBN 0-12-220530-8. £26, \$29.

Solid-State Sciences, Vol. 31. Statistical Physics II: Nonequilibrium Statistical Mechanics. Springer-Verlag: 1985. Pp. 279. ISBN 3-540-11461-0/0-387-11461-0. DM 98.

Spherical Harmonics and Tensors for Classical Field Theory. By M.N. JONES. Research Studies Press: 1985. Pp. 230. ISBN 0-86380-028-9. £28.90.

Springer Series in Optical Sciences: Integrated Optics. H.P. NOLTING and R. ULRICH (ed.). Springer-Verlag: 1985. Pp. 242. ISBN 3-540-15537-6/0-387-15537-6. DM 62.

Tetrahedrally-Bonded Amorphous Semiconductors. DAVID ADLER and HELLMUT FRITZSCHE (eds). Plenum: 1985. Pp. 564. ISBN 0-306-42076-7. Np.

Theory of Heavy Fermions and Valence Fluctuations. Springer Series in Solid-State Sciences 62. T. KASUYA and T. SASO (eds). Springer-Verlag: 1985. Pp. 287. Hbk ISBN 3-540-15922-3. DM 75.

The Theory of Thermodynamics. By J.R. WALDRAM. Cambridge University Press: 1985. Pp. 336. ISBN 0-521-24575-3. Np.

Chemistry

Advances in Organometallic Chemistry, Vol. 24. Academic: 1985. Pp. 470. ISBN 0-12-031124-0. \$79.50, £69.

Adventures with Atoms and Molecules: Chemistry Experiments for Young People. By ROBERT C. MEBANE and THOMAS R. RYBOLD. Enslow: 1985. Pp. 82. ISBN 0-89490-120-6. £10.95.

Aspects of Chemical Evolution. By G. NICOLIS. Wiley: 1985. Pp. 286. ISBN 0-471-88405-7. £52.25.

Handbook of Aqueous Electrolyte Solutions: Physical Properties, Estimation and Correlation Methods. By A.L. HORVATH. Ellis Horwood: 1985. Pp. 631. Hbk ISBN 0-85312-894-4. £69.50.

Homogeneous Catalysis with Compounds of Rhodium and Iridium. By RONALD S. DICKSON. D. Reidel: 1985. Pp. 278. Hbk ISBN 90-277-1880-6. Dfl. 135, £37.50, \$44.50.

Improved Hollow Cathode Lamps for Atomic Spectroscopy. SERGIO CAROLI (ed.) Ellis Horwood: 1985. Pp. 232. ISBN 0-85312-707-7. £35.

Instrumental Methods in Electrochemistry. SOUTHAMPTON ELECTROCHEMISTRY GROUP. Ellis Horwood: 1985. Pp. 443. ISBN 0-85312-875-8. \$49.50.

Molecular Spectroscopy: Modern Research, Vol. III. K. NARAHARI RAO (ed.). Academic: 1985. Pp. 452. ISBN 0-12-580643-4. \$85, £85.

Structural Chemistry of Silicates: Structure, Bonding and Classification. By FRIEDRICH LIEBAU. Springer-Verlag: 1985. Pp. 347. Hbk ISBN 3-540-13747-5. DM 163.

Technology

Dictionary of Robotics. By HARRY WALDMAN. Collier Macmillan: 1985. Pp. 303. Hbk ISBN 0-02-948530-4. £35.

Horticultural Engineering Technology: Field Machinery. By R. BALLS. Macmillan, London: 1985. Pp. 216. Pbk ISBN 0-333-36434-1. Pbk £12.50.

The Manufacture of Soaps, Other Detergents and Glycerine. By EDGAR WOOLLATT. Ellis Horwood, Chichester, UK: 1985. Pp. 473. Hbk ISBN 0-85312-567-8. £55.

Microcircuit Engineering 84. A. HEUBERGER and H. BENEKING (eds). Academic: 1985. Pp. 563. ISBN 0-12-345750-5. \$35, £29.

Process Engineering Developments: The Subject Groups Symposium. Multi-Stream 85. INSTITUTION OF CHEMICAL ENGINEERS. Pergamon Press: 1985. Pp. 398. ISBN 0-08-031445-7. \$27, £22.50.

Radioisotope Techniques for Problem-Solving in Industrial Process Plants. J.S. CHARLTON (ed.). Leonard Hill: 1986. Pp. 320. Hbk ISBN 0-249-44171-3. £35.

Research Techniques in Nondestructive Testing, Vol. VIII. R.S. SHARPE (ed.) Academic: 1985. Pp. 479. ISBN 0-12-639058-4. \$99, £90.

Rock Grouting, With Emphasis on Dam Sites. By F.-K. EWERT. Springer-Verlag: 1985. Pp. 420. Hbk ISBN 3-540-15252-0. DM 175.

Satellite Broadcasting Systems: Planning and Design. By J.N. SLATER and L.A. TRINOGGA. Ellis Horwood: 1985. Pp. 168. Hbk ISBN 0-85312-864-2. £18.50.

Computer Science

How to Write a Usable User Manual. By EDMOND H. WEISS. ISI Press: 1985. Pp. 197. Pbk ISBN 0-89495-052-5. Pbk \$14.95.

Interface Fundamentals in Microprocessor-Controlled Systems. By CHRIS J. GEORGOPOULOS. D. Reidel: 1985. Pp. 364. Hbk ISBN 90-277-2127-0. Dfl. 90, £24.95, \$32.

Earth Sciences

Discours Sur les Revolutions de la Surface du Globe. By GEORGES CUVIER. Christian Bourgeois Editeur: 1985. Pp. 335. Pbk ISBN 2-267-00424-0. Np.

Evolution of Archean Supracrustal Sequences. L.D. AYRES, P.C. THURSTON, K.D. CARD and W. WEBER (eds). Geological Association of Canada: 1985. Pp. 380. ISBN 0-919216-24-2. \$45.

Geological Factors and the Evolution of Plants. BRUCE H. TIFFNEY (ed.) Yale University Press: 1985. Pp. 294. Hbk ISBN 0-300-03304-4. £25.

Geology and Mineral Resources of West Africa. J.B. WRIGHT (ed.). Allen & Unwin: 1985. Pp. 187. ISBN 0-04-556001-3. £30.

Geology for Civil Engineers, 2nd Edn. By A.C. MCLEAN and C.D. GRIBBLE. Allen & Unwin:1985. Pp.314. ISBN 0-04-624006-3. Hbk £20; pbk £9.95.

Geology in Petroleum Production. By A.J. DIKERS. Elsevier:1985. Pp.240. Hbk ISBN 0-444-42450-4. Dfl. 120, \$44.50.

Historical Events and People in Geosciences. WILFRIED SCHRÖDER (ed.). Verlag Peter Lang, Postfach 277, CH-3000, Bern 15, Switzerland:1985. Pp.220. Pbk ISBN 3-8204-7472-2. Pbk SwFr. 48, \$20.85.

Ingenieur-geologische Probleme im Grenzbereich zwischen Lockern und Festgesteinen. KARL-HEINRICH HEITFELD (ed.). Springer-Verlag:1985. Pp. 695. Pbk ISBN 3-540-15366-7. DM 148.

Phytogeomorphology. By J.A. HOWARD and C.W. MITCHELL. Wiley:1985. Pp.222. Hbk ISBN 0-471-09914-7. £46.

Soil Erosion and Conservation. S.A. EL-SWAIFY, W.C. MOLDENHAUER and ANDREW LO (eds). Soil Conservation Society of America:1985. Pp.793. ISBN 0-935734-11-2. \$35.

South African Ocean Colour and Upwelling Experiment. L.V. SHANNON (ed.). Sea Fisheries Research Institute, Capetown, South Africa:1985. Pp.270. Hbk ISBN 0-621-07386-5. Np.

Tectonic Geomorphology. M. MORISAWA and J.T. HACK (eds). Allen & Unwin:1985. Pp.390. ISBN 0-04-551098-9. £25.

Turbulence in the Ocean. By A.S. MONIN and R.V. OZMIDOV. D. Reidel:1985. Pp.247. Hbk ISBN 90-277-1735-4. Np.

A Stratigraphical Index of Conodonts. A.C. HIGGINS and R.L. AUSTIN (eds). Ellis Horwood, Chichester, UK:1985. Pp.263. Hbk ISBN 0-85312-641-0. £27.50.

Environmental Sciences

Acid Rain: Economic Assessment. PAULETTE MANDELBAUM (ed.). Plenum:1985. Pp.287. Hbk ISBN 0-306-42102-X. Np.

Air Pollutants Effects on Forest Ecosystems. The Acid Rain Foundation:1985. Pp.439. \$45.

The Assessment and Control of Major Hazards. INSTITUTION OF CHEMICAL ENGINEERS. Pergamon:1985. Pp.454. ISBN 0-08-031444-9. £30, \$36.

Determination of Organic Substances in Water, Vol. 2. By T.R. CROMPTON. Wiley:1985. Pp.518. ISBN 0-471-90469-4. £49.95.

Environmental Policies: An International Review. CHRIS C. PARK (ed.). Croom Helm, Kent, UK:1985. Pp.315. Hbk ISBN 0-7099-2062-8. £22.50.

Mutagenicity Testing in Environmental Pollution Control. F.K. ZIMMERMANN and R.E. TAYLOR-MAYER (eds). Wiley:1985. Pp.195. ISBN 0-85312-681-X. £27.50.

Regulating Industrial Risks: Science, Hazards and Public Protection. HARRY OTWAY and MALCOLM PELTU (eds). The Butterworth Group, Kent, UK:1985. Pp.181. Hbk ISBN 0-408-00740-0. £25.

Selected Methods of Trace Metal Analysis. Biological and Environmental Samples. By JON C. VAN LOON. Wiley:1985. Pp.357. Hbk ISBN 0-471-89634-9. £56.25.

Biological Sciences

Aphasia and Brain Organization. By IVAR REIN-VANG. Plenum:1985. Pp.195. ISBN 0-306-41975-0. Np.

Arbrégé de Zoologie. 2. Vertébrés. By P.P. GRAS-SÉ. Masson:1985. Pp.184. Pbk ISBN 2-225-80576-8. F 100.

Atherosclerosis Reviews. Arachidonic Acid Metabolites. Vol. 13. RUTH JOHNSON HEGYELI (ed.). Raven Press:1985. Pp.182. Hbk ISBN 0-88167-131-2. \$49.50.

Behaviour and Pathology of Aging in Rhesus Monkeys. ROGER D. DAVIS and CHARLES W. LEATHERS (eds). Alan R. Liss:1985. Pp.380. Hbk ISBN 0-8451-3407-8. £67.00.

Biology of Invertebrate and Lower Vertebrate Col-lagens. NATO ASI Series. A. BAIRATI and R. GAR-RONE (eds). Plenum:1985. Pp.583. Hbk ISBN 0-306-42078-3. Np.

Biology of the Reptilia. CARL GANS, FRANK BILLET and PAUL F.A. MADERSON (eds). Wiley:1985. Pp.763. ISBN 0-471-81358-3. £86.90.

Body and Self: Elements of Human Biology, Behaviour, and Health. By GEORGE BLOCH. W.H. Freeman:1985. Pp.313. ISBN 0-86576-041-1. £17.95.

Calcium and Cell Physiology. DIETER MARME (ed.). Springer-Verlag:1985. Pp.390. ISBN 3-540-13841-2/0-387-13841-2. DM 128.

Cancer Biology: Readings from Scientific American. Introductions by ERROL C. FREIDBERG. W.H. Freeman:1985. Pp.156. Pbk ISBN 0-7167-1751-4. Pbk \$12.95.

Cell Transformation. J. CELIS and A. GRAESS-MANN (eds). Plenum:1985. Pp.321. Hbk ISBN 0-306-42082-1. Np.

Coléoptères Scarabaeoidea. By J. BARAUD. Editions Lechevalier, Paris:1985. Pp.651. Hbk ISBN 2-7205-0517-X. Np.

Common Seaweeds of China. C.K. TSENG (ed.). Science Press, Beijing/Kugler, Amsterdam, The Netherlands:1985. Pp.316. ISBN 90-6299-012-6. Dfl. 400.

Control of Leaf Growth. N.R. BAKER, W.J. DAVIES and C.K. ONG (eds). Cambridge University Press:1985. Pp.350. ISBN 0-521-30480-6. £22.50.

Developmental Biology. By SCOTT F. GILBERT. Blackwell:1985. Pp.726. ISBN 0-87893-246-1. Np.

Developmental Biology of Higher Fungi. D. MOORE et al. (eds). Cambridge University Press:1985. Pp.615. ISBN 0-521-30161-0. Np.

Developments in Biological Standardization. R. SPIER, B. GRIFFITHS and W. HENNESSEN (eds). Karger:1985. Pp.546. Pbk ISBN 3-8055-4043-4. Pbk DM 168, \$59.75.

Energetics of the Photosynthesizing Plant Cell. By L.N. BELL. Harwood:1985. Pp.402. ISBN 3-7186-0195-8. \$175.

Evolutionary Relationships Among Rodents: A Multidisciplinary Analysis. W. PATRICK LUCKETT and JEAN-LOUIS HARTENBERGER (eds). Plenum:1985. Pp.721. ISBN 0-306-42061-9. \$110.

Extracellular Matrix: Structure and Function. A. HARI REDDI (ed.). Alan R. Liss:1985. Pp.462. Hbk ISBN 0-8451-2624-5. £73.00.

Fat-Soluble Vitamins: Their Biochemistry and Applications. ANTHONY T. DIPLOCK (ed.). Heinemann:1985. Pp.319. ISBN 0-434-90312-4. £22.

Fish Immunology. MARGARET J. MANNING and MARY F. TATNER (eds). Academic:1985. Pp.374. Hbk ISBN 0-12-469230-3. £27.00, \$32.50.

Flying Squirrels: Gliders in the Dark. By NANCY WELLS-GOSLING. Smithsonian Institution Press:1985. Pp.128. ISBN 087474-952-2. \$24.95.

Form and Function in Birds, Vol. 3. A.S. KING and J. McLELLAND (eds). Academic:1985. Pp.522. Hbk ISBN 0-12-407503-7. £90.00, \$99.50.

Frankia and Actinorhizal Plants. Developments in Plant and Soil Sciences, Vol. 18. M. LALONDE, C. CAMIRE and J.O. DAWSON (eds). Martinus Nijhoff/Dr W. Junk:1985. Pp.208. Hbk ISBN 90-247-3214-X. Dfl. 120, £33.25, \$45.00.

From Laurel Hill to Siler's Bog. The Walking Adventures of a Naturalist. By JOHN K. TERRES. Algonquin Books of Chapel Hill, PO Box 2225, Chapel Hill, North Carolina, USA:1985. Pp.227. Pbk ISBN 0-912697-26-1. \$12.95.

Frontiers in Histamine Research: A Tribute to Heinz Schild. C.R. GANELLIN and J.C. SCHWARTZ (eds). Pergamon:1985. Pp.442. ISBN 0-08-031989-0. \$80.

Frontiers in Physiological Research. D.G. GAR-LICK and P.I. KORNER (eds). Cambridge University Press:1985. Pp.391. ISBN 0-521-26838-9. £67.50, \$95.

Fungal Dimorphism, with Emphasis on Fungi Pathogenic for Humans. PAUL J. SZANISZLO (ed.). Plenum:1985. Pp.395. Hbk ISBN 0-306-42020-1. Np.

Hormonal Regulation of Plant Growth and Development. S.S. PUROHIT (ed.). Martinus Nijhoff:1985. Pp.412. ISBN 90-247-3198-4. Dfl. 200, \$69.50, £55.50.

Hormones and Cell Regulation: European Symposium, Vol. 9. J.E. DUMONT, B. HAMPRECHT and J. NUNEZ (eds). Elsevier:1985. Pp.430. ISBN 0-7204-0657-9. \$89.

The Illustrated Encyclopedia of Dinosaurs. By Dr DAVID NORMAN. Salamander Books, London, UK:1985. Pp.208. Hbk ISBN 0-86101-225-9. £9.95.

The Illustrated Tarka the Otter. By HENRY WILLIAMSON. Webb & Bower:1985. Pp.207. £12.95.

Immunological Methods, Vol. 3. IVAN LEFKOVITS and BENVENUTO PERNIS (eds). Academic:1985. Pp.477. Hbk ISBN 0-12-442703-0. £48.50, \$54.50.

Institute of Terrestrial Ecology: Annual Report 1984. Natural Environment Research Council:1985. Pp.188. ISBN 0-904282-84-8. Np.

International Review of Cytology, Vol. 93. Genome Evolution in Prokaryotes and Eukaryotes. D.C. REANNEY and P. CHAMBER. Academic:1985. Pp.367. ISBN 0-12-364493-3. \$49.50, £49.50.

Islands in the Bush: A Natural History of the Kora National Reserve, Kenya. By MALCOLM COE. George Philip:1985. Pp.240. ISBN 0-540-01086-3. £14.95.

The Living Isles: A Natural History of Britain and Ireland. By PETER CRAWFORD. BBC Books:1985. Pp.320. Hbk ISBN 0-563-20369-2. £14.95.

Longman Illustrated Dictionary of Biology: Living Organisms in All Forms Explained and Illustrated. By NEIL CURTIS. Longman:1985. Pp.256. ISBN 0-582-89255-4. £3.95.

Longman World Guide to Mammals. PHILIP WHITFIELD (ed.). Longman:1985. Pp.198. ISBN 0-582-89211-2. £12.95.

Microenvironments in the Lymphoid System. G.G.B. KLAUS (ed.). Plenum:1985. Pp.1080. ISBN 0-306-41917-3. Np.

Multiple Choice Questions in Neurophysiology with Answers and Explanatory Comments. By LYNN BINDMAN and PETER ELLAWAY. Edward Arnold:1985. Pp.93. Pbk ISBN 0-7131-4474-2. Pbk £3.95.

The Natural History of Otters. By PAUL CHANIN. Croom Helm:1985. Pp.179. ISBN 0-7099-3460-2. £7.95.

The Nature Watchers: Exploring Wildlife with the Experts. By R. BROWN and J. PETTIFER. Collins:1985. Pp.240. Hbk ISBN 0-00-219149-0. £12.00.

The Neuropsychology of Individual Differences: A Developmental Perspective. LAWRENCE C. HART-LAGE and CATHY F. TELZROW (eds). Plenum:1985. Pp.329. ISBN 0-306-41986-6. \$35.

The New Theory of Respiratory Dynamics. By A. GONZALEZ-BOGEN. Universidad Central de Venezuela, Epícoras de la Biblioteca, Caracas:1985. Pp.141. Pbk ISBN 3125901. Np.

Notes of an Anatomist. By F. GONZALEZ-CRUSSI. Harcourt Brace Jovanovich:1985. Pp.144. ISBN 0-15-167285-7. \$12.95.

Nuclear Envelope Structure and RNA Maturation. EDWARD A. SMUCKLER and GARY A. CLAWSON (eds). Alan R. Liss:1985. Pp.694. Hbk ISBN 0-8451-2625-3. £75.00.

The Origins and Relationships of Lower Invertebrates. S. CONWAY MORRIS, J.D. GEORGE, R. GIBSON and H.M. PLATT (eds). Oxford University Press:1985. Pp.397. ISBN 0-19-857181-X. £45.

The Oxford Dictionary of Natural History. MICHAEL ALLABY (ed.), DAVID ATTENBOROUGH (forward). Oxford University Press:1985. Pp.688. Hbk ISBN 0-19-217720-6. £20.00.

Oxford Reviews in Reproductive Biology, Vol. 7. J.R. CLARKE (ed.). Clarendon Press:1985. Pp.409. ISBN 0-19-857540-8. Np.

People of the Desert and Sea. Ethnobotany of the Seri Indians. By RICHARD FELGER and MARY BECK MOSER. University of Arizona Press:1985. Pp.435. \$65.

The Physiological Development of Fetus and New-born. C.T. JONES and P.W. NATHANIELSZ (eds). Academic:1985. Pp.837. ISBN 0-12-389080-2. \$54, £49.

Proteins of Iron Storage and Transport. G. SPIK, J. MONTREUIL, R.R. CRICHTON and J. MAZURIER (eds). Elsevier:1985. Pp.380. Hbk ISBN 0-444-80722-5. Dfl. 198, \$73.25.